

ESCUELA POLITÉCNICA NACIONAL

FACULTAD DE INGENIERÍA DE SISTEMAS

**IMPLEMENTACIÓN DE UN DATA MART PARA EL RAMO DE
VEHÍCULOS EN UNA EMPRESA ASEGURADORA**

**PROYECTO PREVIO A LA OBTENCIÓN DEL TÍTULO DE INGENIERO
EN SISTEMAS INFORMÁTICOS Y DE COMPUTACIÓN**

CARLOS MAURICIO MURILLO ROSERO

Email: mauricio2k4@gmail.com

DIRECTOR: ING. FRANCISCO VILLAVICENCIO

Email: fvillavi@yahoo.com

Quito, Diciembre 2008

DECLARACIÓN

Yo Carlos Mauricio Murillo Rosero, declaro bajo juramento que el trabajo aquí descrito es de mi autoría; que no ha sido previamente presentada para ningún grado o calificación profesional; y, que he consultado las referencias bibliográficas que se incluyen en este documento.

A través de la presente declaración cedo mis derechos de propiedad intelectual correspondientes a este trabajo, a la Escuela Politécnica Nacional, según lo establecido por la Ley de Propiedad Intelectual, por su Reglamento y por la normatividad institucional vigente.

Carlos Mauricio Murillo Rosero

CERTIFICACIÓN

Certifico que el presente trabajo fue desarrollado por Carlos Mauricio Murillo Rosero, bajo mi supervisión.

Ing. Francisco Villavicencio

DIRECTOR DE PROYECTO

DEDICATORIA

Dedico este proyecto de titulación a mis padres, Gustavo y Zoily, quienes han sido los pilares de toda mi vida. Ellos, quienes me han sabido aconsejar y guiar por buenos senderos. Ellos, quienes me han apoyado en los buenos y malos momentos. Ellos, quienes me han apoyado incondicionalmente en todo ámbito para que pueda cumplir mis metas y mis más grandes sueños.

Dedico este proyecto de titulación a mis amigos que me acompañaron durante mi carrera universitaria. Mis amigos de combate, con los cuáles tuvimos excelentes y malos ratos, supimos aprovechar al máximo todo lo aprendido y pudimos levantarnos con mucha más fuerza cuando nos caímos. Gracias a ellos por sus ideas, apoyo incondicional y en especial por su amistad. Para mis grandes amigos, Guillermo, Marco y José.

Dedico este proyecto de titulación al resto de mi familia y a mis hermanos quienes estuvieron pendientes por que termine con ésta meta, apoyándome y dándome ánimos día a día. Dedico este trabajo en especial a mi abuelita Lucy, mi tía Hilde, y mi hermana Carolina, mujeres a quienes amo con todo mi corazón.

Dedico este proyecto de titulación a toda la comunidad politécnica, a mis profesores, compañeros de aula y mis jefes de la empresa en la que actualmente laboro. Todos ellos que me compartieron su conocimiento, me enseñaron a ser profesional y a la vez a ser una excelente persona.

Gracias a todos ustedes.

Mauricio Murillo

INTRODUCCION

El crecimiento de los sistemas de información para la gestión de diversas áreas que maneja una Compañía, ha conducido a que las empresas se vuelvan ricas en datos pero pobres en información. El presente proyecto de titulación tiene como objetivo demostrar cómo la inteligencia de negocios, Business Intelligence en inglés, ayuda a las compañías a explotar ese activo tan importante, sus datos, para transformarlos en información útil que apoye al proceso de toma de decisiones estratégicas, tácticas y operativas.

En el presente proyecto de titulación, de título, Implementación de un Data Mart para el ramo de vehículos en una empresa aseguradora, se ha tomado como caso de estudio el negocio asegurador, construyéndose un Data Mart de análisis de producción y siniestros para pólizas de vehículos. Se ha hecho énfasis en las arquitecturas de inteligencia de negocios propuestas por los padres del Data Warehousing, Bill Inmon y Ralph Kimball, y se ha enfocado el proyecto hacia la utilización de una plataforma de software para inteligencia de negocios que soporte cualquiera de las arquitecturas propuestas.

La plataforma de inteligencia de negocios seleccionada es BusinessObjects Enterprise XI Release 2. BusinessObjects se adapta a cualquier arquitectura de inteligencia de negocios y apoya a todos los procesos de un sistema de Data Warehouse que van desde la obtención de datos de entrada hasta la entrega de información de salida.

El proyecto de titulación se lo ha estructurado en cinco capítulos. El Capítulo I hace énfasis en las arquitecturas de inteligencia de negocios propuestas por Inmon y Kimball. Se explica y describe la fábrica de información corporativa (CIF, Corporate Information Factory) propuesta por Inmon y la arquitectura multidimensional (MD, Multidimensional) propuesta por Kimball. Se realiza una comparación entre las arquitecturas de inteligencia de negocios y se explica porqué se decidió implementar la arquitectura Multidimensional en el proyecto de titulación. Se hace énfasis también en los conceptos de inteligencia de negocios, Data Warehousing, las arquitecturas de flujos de datos para los sistemas de Data

Warehouse y se definen las metodologías de desarrollo para la implementación de un sistema de Data Warehouse. Además, se explican las razones por las cuáles se seleccionó la metodología en cascada para el desarrollo del proyecto de titulación.

En el Capítulo II se realiza un análisis del caso de estudio, haciéndose énfasis en la explicación de los términos y definiciones del negocio asegurador. Se define el problema a solucionar y los términos correspondientes de la empresa de seguros, el ramo de vehículos de una empresa aseguradora y los medios técnicos para la gestión de una empresa de seguros. Éste capítulo explica los términos del negocio asegurador para que sea de fácil comprensión la implementación del Data Mart que resuelve el problema definido.

El Capítulo III, constituye el capítulo más importante del proyecto de titulación. En éste capítulo se explican todos los pasos para la construcción del Data Mart. Se define la arquitectura de flujo de datos adoptada y cómo las herramientas de la plataforma de inteligencia de negocios BusinessObjects apoyan a cada etapa del proceso de construcción del Data Mart. En éste capítulo se definen los requerimientos funcionales y no funcionales, se explica el diseño de la base de datos OLTP para seguros de vehículos y se diseña el modelo dimensional que cumple con los requerimientos funcionales del Data Mart. En éste capítulo también se explica cómo se construyeron los procesos de extracción, transformación y carga (ETL, Extract, Transform and Load) utilizando la herramienta de integración de datos BusinessObjects Data Integrator, perteneciente a la plataforma de inteligencia de negocios BusinessObjects. Además, se explica como se integra la herramienta ETL con la herramienta de entrega de información a usuarios finales BusinessObjects InfoView. Finalmente, se explican dos mecanismos para realizar las pruebas de consistencia de datos para asegurarse que la información generada por el sistema OLTP sea consistente con la información generada por el Data Mart.

En el Capítulo IV se obtiene información del Data Mart construido y se la presenta a través de informes y un tablero de control. En éste capítulo se utiliza la herramienta de entrega de información BusinessObjects InfoView, la herramienta de reporte BusinessObjects Web Intelligence y la herramienta de gestión del

rendimiento BusinessObjects Performance Management. Los informes generados presentan información de análisis con los indicadores de gestión del negocio de seguros más importantes para producción y siniestros de pólizas de vehículos, a través de los informes generados se muestra la actividad de análisis más básica de inteligencia de negocios, el reporte. El tablero de mando muestra los indicadores del negocio asegurador más relevantes usándose indicadores clave de rendimiento (KPI, Key Performance Indicator). A través de analíticas se presentan los indicadores clave de rendimiento para la gestión del ramo de vehículos; mostrándose la gestión del rendimiento (BPM, Business Performance Management) como otra actividad de inteligencia de negocios aplicada al Data Mart.

En el Capítulo V, se describen las conclusiones y recomendaciones del proyecto de titulación, haciéndose énfasis en la utilización de una plataforma de inteligencia de negocios que apoye a todos los procesos de un sistema de Data Warehouse e incentivando a futuros ingenieros a tomar el presente proyecto de titulación como base para aplicar la minería de datos, actividad de análisis de inteligencia de negocios poco explotada en el medio, desafiante pero a la vez muy interesante para aplicarla en el negocio asegurador.

CONTENIDO

INTRODUCCIÓN

CAPÍTULO I	1
1 MARCO TEÓRICO	1
1.1 INTELIGENCIA DE NEGOCIOS	1
1.1.1 INTRODUCCIÓN A LA INTELIGENCIA DE NEGOCIOS	1
1.1.2 DEFINICIÓN DE INTELIGENCIA DE NEGOCIOS.....	2
1.1.2.1 Reporteo	3
1.1.2.2 OLAP (Online Analytical Processing)	4
1.1.2.3 Minería de datos (Data Mining)	5
1.1.3 ARQUITECTURAS DE INTELIGENCIA DE NEGOCIOS	5
1.1.3.1 Corporate Information Factory (CIF)	6
1.1.3.2 Arquitectura Multidimensional (MD)	9
1.1.3.3 Comparativa entre CIF y la arquitectura multidimensional (MD).....	11
1.1.4 HERRAMIENTAS DE INTELIGENCIA DE NEGOCIOS	14
1.1.4.1 BusinessObjects Enterprise XI Release 2.....	15
1.1.5 TENDENCIAS EN LA INTELIGENCIA DE NEGOCIOS	16
1.1.5.1 BICC, Business Intelligence Competency Center	17
1.2 DATA WAREHOUSING	19
1.2.1 PERSPECTIVA HISTÓRICA.....	19
1.2.2 DEFINICIÓN DE DATA WAREHOUSE.....	20
1.2.2.1 Recuperación de datos.....	21
1.2.2.2 Consolidación de datos.....	21
1.2.2.3 Datos actualizados periódicamente	23
1.2.2.4 Almacén de datos	23
1.2.2.5 Historia.....	24
1.2.2.6 Consultas.....	24
1.2.3 DEFINICIÓN DE DATA MART.....	24
1.2.4 ARQUITECTURA DE UN DATA WAREHOUSE.....	26
1.2.4.1 Arquitectura de flujo de datos	27
1.2.4.1.1 Arquitectura de flujo de datos general	28
1.2.4.1.2 Arquitectura simple con un almacén de datos dimensional (Simple DDS).....	30
1.2.4.1.3 Arquitectura con un almacén de datos normalizado más un almacén de datos dimensional (NDS + DDS).....	31
1.2.4.1.4 Arquitectura con un almacén de datos operacional más un almacén de datos dimensional (ODS + DDS)	32
1.2.4.2 Arquitectura del sistema.....	33
1.3 METODOLOGÍA DE DESARROLLO	34
1.3.1 METODOLOGÍA EN CASCADA	35
1.3.1.1 Estudio de factibilidad.....	36
1.3.1.2 Requerimientos	36

1.3.1.3	Arquitectura	36
1.3.1.4	Diseño	37
1.3.1.5	Desarrollo.....	37
1.3.1.6	Pruebas.....	37
1.3.1.7	Implantación.....	37
1.3.1.8	Configuración de la infraestructura	37
1.3.1.9	Gestión del proyecto	37
1.3.2	<i>METODOLOGÍA ITERATIVA</i>	38
CAPÍTULO II.....		40
2	ANÁLISIS DEL CASO DE ESTUDIO.....	40
2.1	DESCRIPCIÓN DEL PROBLEMA.....	40
2.1.1	<i>OBJETIVOS</i>	42
2.1.2	<i>ALCANCE</i>	42
2.2	LA EMPRESA DE SEGUROS	42
2.2.1	<i>DEFINICIÓN</i>	42
2.2.2	<i>EL RIESGO</i>	43
2.2.3	<i>EL SEGURO</i>	44
2.2.3.1	El Contrato de seguro.....	45
2.2.3.1.1	Elementos materiales	46
2.2.3.1.2	Elementos personales.....	47
2.2.3.1.3	Elementos formales.....	48
2.2.4	<i>EL REASEGURO</i>	50
2.2.5	<i>CLASES DE SEGUROS</i>	50
2.2.6	<i>ORGANIZACIÓN DE UNA EMPRESA DE SEGUROS</i>	51
2.2.6.1	Organización funcional de una empresa de seguros.....	53
2.2.6.1.1	Según el grado de centralización.....	53
2.2.6.1.2	Según la ubicación geográfica de la estructura administrativa.....	53
2.2.6.1.3	Según el grado de agrupación de servicios de idéntica naturaleza	53
2.2.6.2	Órganos de dirección y departamentos de una empresa de seguros	54
2.2.6.3	Proyección de una empresa de seguros hacia el futuro.....	56
2.2.6.3.1	Política comercial	56
2.2.6.3.2	Establecimiento de objetivos	56
2.2.6.3.3	Función comercial y estrategia empresarial	57
2.3	EL RAMO DE VEHÍCULOS Y LOS MEDIOS TÉCNICOS PARA SU GESTIÓN.....	58
2.3.1	<i>LOS RAMOS Y SU CLASIFICACIÓN</i>	58
2.3.2	<i>EL SEGURO DE VEHÍCULOS</i>	59
2.3.2.1	Definiciones Generales	60
2.3.2.1.1	Automotor.....	60
2.3.2.1.2	Tipos de vehículos	60
2.3.2.2	Póliza de vehículos.....	60
2.3.2.2.1	Componentes de la póliza de vehículos	61
2.3.2.2.2	Principales coberturas de la póliza de vehículos	62
2.3.2.2.3	Principales exclusiones de la póliza de vehículos	63

2.3.3	<i>LOS MEDIOS TÉCNICOS DEL SEGURO</i>	64
2.3.3.1	Resultado técnico	64
2.3.3.2	Índice de frecuencia	65
2.3.3.3	Índice de intensidad	65
2.3.3.4	Índice de siniestralidad	66
2.3.3.5	Siniestralidad	66
2.3.3.6	Provisiones técnicas o reservas	66
2.3.3.7	Otros indicadores de análisis	66
CAPÍTULO III		68
3	DESARROLLO DEL DATA MART	68
3.1	PLANEACIÓN	69
3.1.1	<i>HERRAMIENTAS</i>	69
3.1.2	<i>RECURSOS</i>	72
3.1.3	<i>ACTIVIDADES A REALIZAR</i>	73
3.2	DEFINICIÓN DE REQUERIMIENTOS	75
3.2.1	<i>REQUERIMIENTOS FUNCIONALES</i>	75
3.2.2	<i>REQUERIMIENTOS NO FUNCIONALES</i>	78
3.3	DISEÑO	79
3.3.1	<i>BASE DE DATOS DE SEGUROS DE VEHÍCULOS</i>	80
3.3.1.1	Modelo de datos lógico	80
3.3.1.2	Diccionario de datos	82
3.3.1.2.1	Tabla AGENCIA	82
3.3.2	<i>DATA MART PARA ANÁLISIS DE SEGUROS DE VEHÍCULOS</i>	82
3.3.2.1	Modelo dimensional de datos	82
3.3.2.2	Diccionario de datos	85
3.3.2.2.1	Dimensión DIM_AGENCIA	85
3.3.2.2.2	Dimensión DIM_AGENTE_SEGUROS	86
3.3.2.2.3	Dimensión DIM_ASEGURADO	86
3.3.2.2.4	Dimensión DIM_BENEFICIARIO	87
3.3.2.2.5	Dimensión DIM_CLIENTE	87
3.3.2.2.6	Dimensión DIM_COBERTURA	88
3.3.2.2.7	Dimensión DIM_ENDOSO	88
3.3.2.2.8	Dimensión DIM_EXTRA_ASEGURADO	89
3.3.2.2.9	Dimensión DIM_FECHA	89
3.3.2.2.10	Dimensión DIM_UNIDAD_NEGOCIO	90
3.3.2.2.11	Dimensión DIM_VEHICULO	90
3.3.2.2.12	Tabla de Hechos FACT_PRODUCION	92
3.3.2.2.13	Tabla de Hechos FACT_SINIESTROS	93
3.4	CONSTRUCCIÓN DEL DATA MART	95
3.4.1	<i>ARQUITECTURA DE FLUJO DE DATOS</i>	95
3.4.2	<i>HERRAMIENTA BUSINESSOBJECTS DATA INTEGRATOR</i>	96
3.4.2.1	Introducción	96
3.4.2.2	Arquitectura de Data Integrator	99

3.4.3	<i>IMPLEMENTACIÓN DE FLUJOS DE CARGA</i>	100
3.4.3.1	Creación de almacenes de datos (Datastores)	101
3.4.3.2	JOB_CARGA_UNICA	103
3.4.3.3	JOB_CARGA_DATOS	106
3.4.3.3.1	Flujo de Trabajo WF_SRC_DATA_STAGE	107
3.4.3.3.2	Flujo de Trabajo WF_STAGE_DIM_DATA	132
3.4.4	<i>CREACIÓN DEL UNIVERSO DE DATOS</i>	149
3.4.5	<i>PRUEBAS</i>	154
3.4.5.1	Cuadre de datos de producción	155
3.4.5.2	Cuadre de datos de siniestros	160
CAPÍTULO IV		163
4	APLICACIÓN DEL DATA MART AL CASO DE ESTUDIO	163
4.1	DEFINICIÓN DE LOS INDICADORES DE GESTIÓN	163
4.2	REPORTES CON LOS INDICADORES DE GESTIÓN	167
4.2.1	<i>REPORTE DE RESULTADO TÉCNICO</i>	169
4.2.2	<i>REPORTE DE INDICADORES DE ANÁLISIS GLOBALES</i>	172
4.2.3	<i>REPORTE DE ANÁLISIS DE TASAS</i>	175
4.2.4	<i>REPORTE DE ANÁLISIS DE RECLAMOS</i>	178
4.3	TABLERO DE MANDO INTEGRAL CON LOS INDICADORES DE GESTIÓN MÁS RELEVANTES	181
CAPÍTULO V		187
5	CONCLUSIONES Y RECOMENDACIONES	187
5.1	CONCLUSIONES	187
5.2	RECOMENDACIONES	189
BIBLIOGRAFÍA		191
ANEXOS		194

INDICE DE TABLAS

TABLA 1.1 TABLA COMPARATIVA ENTRE CIF Y MD.....	13
TABLA 3.1 CARACTERÍSTICAS DE HARDWARE Y SOFTWARE DEL EQUIPO QUE SE USARÁ PARA LA CONSTRUCCIÓN DEL DATA MART PARA EL RAMO DE VEHÍCULOS	73
TABLA 3.2 TABLA DE REQUERIMIENTOS FUNCIONALES DEL DATA MART.....	77
TABLA 3.3 TABLA DE REQUERIMIENTOS NO FUNCIONALES DEL DATA MART	78
TABLA 3.4 TABLA AGENCIA	82
TABLA 3.39 DIMENSIÓN DIM_AGENCIA	85
TABLA 3.40 DIMENSIÓN DIM_AGENTE_SEGUROS	86
TABLA 3.41 DIMENSIÓN DIM_ASEGURADO.....	86
TABLA 3.42 DIMENSIÓN DIM_BENEFICIARIO.....	87
TABLA 3.43 DIMENSIÓN DIM_CLIENTE.....	87
TABLA 3.44 DIMENSIÓN DIM_COBERTURA	88
TABLA 3.45 DIMENSIÓN DIM_ENDOSO	89
TABLA 3.46 DIMENSIÓN DIM_EXTRA_ASEGURADO.....	89
TABLA 3.47 DIMENSIÓN DIM_FECHA	90
TABLA 3.48 DIMENSIÓN DIM_UNIDAD_NEGOCIO.....	90
TABLA 3.49 DIMENSIÓN DIM_VEHICULO	91
TABLA 3.50 TABLA DE HECHOS FACT_PRODUCION	93
TABLA 3.51 TABLA DE HECHOS FACT_SINIESTROS.....	95
TABLA 3.52 CAMPOS GENERADOS POR EL QUERY DEL FLUJO DE DATOS DF_LOAD_DIM_FECHA	105
TABLA 3.53 CAMPOS GENERADOS EN EL QUERY DEL FLUJO DE DATOS DF_LOAD_VEHICULO_003.....	116
TABLA 3.54 TRANSFORMADORES DEL DISEÑADOR DE DATA INTEGRATOR PARA APLICAR LA TÉCNICA DE CONTROL DE CAMBIOS SCD TIPO 2.....	134
TABLA 3.55 DATOS OBTENIDOS CON EL SCRIPT SQL PARA CUADRE DE DATOS ENTRE EL OLTP Y EL DATA MART.....	159
TABLA 3.56 DATOS DE SINIESTROS PARA ENERO 2006 GENERADOS A TRAVÉS DE WEB INTELLIGENCE.....	162
TABLA 4.1 INDICADORES DE GESTIÓN DE SEGUROS.....	167
TABLA 4.2 REPORTES DE ANÁLISIS CONSTRUIDOS	168

INDICE DE FIGURAS

FIGURA 1.1 LA FÁBRICA DE INFORMACIÓN CORPORATIVA (CIF) DE BILL INMON [4]	7
FIGURA 1.2 LA ARQUITECTURA MULTIDIMENSIONAL (MD) DE RALPH KIMBALL [4]	10
FIGURA 1.3 COMPARATIVA ENTRE CIF Y MD [4]	12
FIGURA 1.4 LA PLATAFORMA DE BI BUSINESS OBJECTS ENTERPRISE XI RELEASE 2 [7]	15
FIGURA 1.5 ILUSTRACIÓN DE LA DEFINICIÓN DE DATA WAREHOUSE	21
FIGURA 1.6 DATA MART DEPENDIENTE, PROPUESTO POR BILL INMON	25
FIGURA 1.7 DATA MART INDEPENDIENTE, PROPUESTO POR RALPH KIMBALL	26
FIGURA 1.8 ARQUITECTURA DE FLUJO DE DATOS GENERAL PARA UN SISTEMA DE DATA WAREHOUSE [5]	29
FIGURA 1.9 ARQUITECTURA SIMPLE DDS [5]	31
FIGURA 1.10 ARQUITECTURA (NDS + DDS) [5]	32
FIGURA 1.11 ARQUITECTURA (ODS + DDS) [5]	33
FIGURA 1.12 ARQUITECTURA FÍSICA DE UN SISTEMA DE DATA WAREHOUSE [5]	34
FIGURA 1.13 METODOLOGÍA EN CASCADA [5]	36
FIGURA 1.14 METODOLOGÍA ITERATIVA	38
FIGURA 2.1 ORGANIGRAMA IDEAL DE UNA EMPRESA DE SEGUROS [6]	52
FIGURA 2.2 ORGANIZACIÓN VERTICAL DE UNA EMPRESA DE SEGUROS [6]	54
FIGURA 2.3 ORGANIZACIÓN HORIZONTAL DE UNA EMPRESA DE SEGUROS [6]	54
FIGURA 2.4 FÓRMULA DE CÁLCULO DEL RESULTADO TÉCNICO	65
FIGURA 2.5 FÓRMULA DE CÁLCULO DEL ÍNDICE DE FRECUENCIA PARA UN AÑO COMPLETO	65
FIGURA 2.6 FÓRMULA DE CÁLCULO DEL ÍNDICE DE FRECUENCIA EN BASE A LA CARTERA DE VEHÍCULOS ASEGURADOS	65
FIGURA 2.7 FÓRMULA DE CÁLCULO DEL ÍNDICE DE INTENSIDAD O ÍNDICE DE SINIESTRO PROMEDIO	65
FIGURA 2.8 FÓRMULA DE CÁLCULO DEL ÍNDICE DE SINIESTRALIDAD	66
FIGURA 2.9 FÓRMULA DE CÁLCULO DEL VALOR ASEGURADO PROMEDIO	67
FIGURA 2.10 FÓRMULA DE CÁLCULO DE LA TASA PROMEDIO DE LA CARTERA	67
FIGURA 2.11 FÓRMULA DE CÁLCULO DEL % DE COMISIONES PAGADAS	67
FIGURA 2.12 FÓRMULA DE CÁLCULO DEL % DE GASTOS ADMINISTRATIVOS	67
FIGURA 3.1 HERRAMIENTAS DE LA PLATAFORMA BOE XI R2 SP3	71
FIGURA 3.2 MODELO DE DATOS LÓGICO PARA SEGUROS DE VEHÍCULOS	80
FIGURA 3.3 MODELO DE DATOS DIMENSIONAL PARA EL DATA MART	83
FIGURA 3.4 ARQUITECTURA DE FLUJO DE CARGA DE DATOS QUE SE IMPLEMENTARÁ EN LA HERRAMIENTA BUSINESSOBJECTS DATA INTEGRATOR	95
FIGURA 3.5 PANTALLA DE ACCESO A DATA INTEGRATOR	96
FIGURA 3.6 PANTALLA PRINCIPAL DE DATA INTEGRATOR DESIGNER	97
FIGURA 3.7 JERARQUÍA DE OBJETOS MANEJADOS POR DATA INTEGRATOR DESIGNER	98
FIGURA 3.8 COMPONENTES ESTÁNDAR DE DATA INTEGRATOR [35]	99
FIGURA 3.9 ESTRUCTURA DEL PROYECTO DM_SEGUROS_VEHICULOS EN DATA INTEGRATOR DESIGNER	101
FIGURA 3.10 PROCESO PARA CREAR ALMACENES DE DATOS EN DATA INTEGRATOR DESIGNER	102
FIGURA 3.11 FLUJO DE DATOS PARA CARGAR LA DIMENSIÓN FECHA (DF_LOAD_DIM_FECHA)	103

FIGURA 3.12 CONSULTA (QUERY) DEL FLUJO DE DATOS DF_LOAD_DIM_FECHA	104
FIGURA 3.13 DIMENSIÓN FECHA (DIM_FECHA) CARGADA DE DATOS	106
FIGURA 3.14 ESTRUCTURA DE LOS FLUJOS DE TRABAJO DEL JOB_CARGA_DATOS.....	106
FIGURA 3.15 FLUJO DE DATOS DF_LOAD_COBERTURA_001	107
FIGURA 3.16 CARGA DE LA TABLA TEMPORAL MAPEO_COBERTURA	108
FIGURA 3.17 FLUJO DE DATOS DF_LOAD_COBERTURA_002.....	108
FIGURA 3.18 FLUJO DE DATOS DF_LOAD_AGENCIA_ENCARGADO	109
FIGURA 3.19 FLUJOS DE DATOS DEL FLUJO DE TRABAJO ETL_ASEGURADO	110
FIGURA 3.20 FLUJOS DE DATOS DEL FLUJO DE TRABAJO ETL_BENEFICIARIO	111
FIGURA 3.21 FLUJO DE DATOS DF_CLEAN_EXTRA_ASEGURADOS_001	112
FIGURA 3.22 FLUJO DE DATOS DF_CLEAN_EXTRA_ASEGURADOS_001	113
FIGURA 3.23 FLUJO DE DATOS DF_LOAD_EXTRA_ASEGURADO_003 Y 004	113
FIGURA 3.24 FLUJO DE DATOS DF_CLEAN_VEHICULO_001	114
FIGURA 3.25 FLUJO DE DATOS DF_CLEAN_VEHICULO_002	115
FIGURA 3.26 FLUJO DE DATOS DF_LOAD_VEHICULO_003.....	116
FIGURA 3.27 FLUJO DE DATOS DF_LOAD_VEHICULO_004.....	117
FIGURA 3.28 FLUJO DE DATOS DF_LOAD_CLIENTE_001	117
FIGURA 3.29 FLUJO DE DATOS DF_LOAD_ENDOSO_001.....	118
FIGURA 3.30 FLUJO DE DATOS DF_LOAD_ENDOSO_002.....	119
FIGURA 3.31 FLUJO DE TRABAJO ETL_PRODUCION_PRIMA_SUSCRITA	119
FIGURA 3.32 FLUJO DE DATOS DF_LOAD_PRIMA_SUSCRITA_001	120
FIGURA 3.33 FLUJO DE DATOS DF_LOAD_PRIMA_SUSCRITA_002	121
FIGURA 3.34 FLUJO DE DATOS DF_LOAD_PRIMA_SUSCRITA_003	122
FIGURA 3.35 FLUJO DE DATOS DF_LOAD_PRIMA_SUSCRITA_004	122
FIGURA 3.36 FLUJO DE DATOS DF_LOAD_PRIMA_SUSCRITA_005	123
FIGURA 3.37 FLUJO DE TRABAJO ETL_PRODUCION_PRIMA_COBRADA	124
FIGURA 3.38 FLUJO DE DATOS DF_LOAD_PRIMA_COBRADA_001	124
FIGURA 3.39 FLUJO DE DATOS DF_LOAD_PRIMA_COBRADA_002	125
FIGURA 3.40 FLUJO DE DATOS DF_LOAD_PRIMA_COBRADA_003	126
FIGURA 3.41 FLUJO DE DATOS DF_LOAD_PRIMA_COBRADA_004	126
FIGURA 3.42 FLUJO DE DATOS DF_LOAD_PRIMA_COBRADA_007	127
FIGURA 3.43 ESTRUCTURA DEL FLUJO DE TRABAJO ETL_SINIESTROS_POLIZAS	128
FIGURA 3.44 FLUJO DE DATOS DF_LOAD_SINIESTROS_BASE_001	129
FIGURA 3.45 FLUJO DE DATOS DF_LOAD_SINIESTROS_BASE_002	130
FIGURA 3.46 FLUJO DE DATOS DF_LOAD_INDEMNIZACIONES	131
FIGURA 3.47 FLUJO DE DATOS DF_LOAD_RESERVAS	132
FIGURA 3.48 FLUJO DE DATOS DF_LOAD_DIM_COBERTURA	133
FIGURA 3.49 FLUJO DE DATOS DF_LOAD_DIM_AGENCIA	135
FIGURA 3.50 FLUJO DE DATOS DF_LOAD_DIM_UNIDAD_NEGOCIO.....	136
FIGURA 3.51 FLUJO DE DATOS DF_LOAD_DIM_ASEGURADO.....	137
FIGURA 3.52 FLUJO DE DATOS DF_LOAD_DIM_BENEFICIARIO	138
FIGURA 3.53 FLUJO DE DATOS DF_LOAD_DIM_EXTRA_ASEGURADO	138

FIGURA 3.54 FLUJO DE DATOS DF_LOAD_DIM_VEHICULO.....	139
FIGURA 3.55 FLUJO DE DATOS DF_LOAD_DIM_AGENTE.....	140
FIGURA 3.56 FLUJO DE DATOS DF_LOAD_DIM_CLIENTE.....	141
FIGURA 3.57 FLUJO DE DATOS DF_LOAD_DIM_ENDOSO.....	142
FIGURA 3.58 FLUJO DE DATOS DF_LOAD_FACT_PRODUCCION_001.....	143
FIGURA 3.59 FLUJO DE DATOS DF_LOAD_FACT_PRODUCCION_002.....	144
FIGURA 3.60 FLUJO DE DATOS DF_LOAD_FACT_PRODUCCION_003.....	144
FIGURA 3.61 FLUJO DE DATOS DF_LOAD_FACT_PRODUCCION.....	145
FIGURA 3.62 FLUJO DE DATOS DF_LOAD_FACT_SINIESTROS_001.....	146
FIGURA 3.63 FLUJO DE DATOS DF_LOAD_FACT_SINIESTROS_002.....	147
FIGURA 3.64 FLUJO DE DATOS DF_LOAD_FACT_SINIESTROS.....	147
FIGURA 3.65 FLUJO DE DATOS DF_LOAD_PRODUCCION_SINIESTROS.....	148
FIGURA 3.66 SELECCIÓN DEL DATASTORE PARA CREAR UN UNIVERSO.....	150
FIGURA 3.67 ASISTENTE PARA LA GENERACIÓN DE UNIVERSOS DE DATOS.....	150
FIGURA 3.68 SELECCIÓN DE ELEMENTOS PARA CREAR EL UNIVERSO.....	151
FIGURA 3.69 CONFIGURACIÓN DE CONEXIÓN Y NOMBRE DEL UNIVERSO.....	151
FIGURA 3.70 RESUMEN DE CONFIGURACIÓN DE CREACIÓN DE UNIVERSOS.....	152
FIGURA 3.71 UNIVERSO DE DATOS DM_VEHIC.....	152
FIGURA 3.72 UNIVERSO DE DATOS DM_VEHIC_PM.....	153
FIGURA 3.73 AUTO CONTEXTO SOBRE LA DIMENSIÓN FECHA PARA CREAR TABLEROS DE MANDO EN BUSINESSOBJECTS PERFORMANCE MANAGEMENT.....	154
FIGURA 3.74 REPORTE DE DETALLE PRODUCCIÓN ENERO 2005 PRIMERA PÁGINA.....	155
FIGURA 3.75 REPORTE DE DETALLE PRODUCCIÓN ENERO 2005 ÚLTIMA PÁGINA.....	156
FIGURA 3.76 INFORME DE WEB INTELLIGENCE GENERADO PARA CUADRAR PRODUCCIÓN DE ENERO 2005 ...	157
FIGURA 3.77 DATOS DE PRODUCCIÓN GENERADOS PARA ENERO 2005 A TRAVÉS DE WEB INTELLIGENCE.....	157
FIGURA 3.78 REPORTE DE PRODUCCIÓN DE ENERO 2005 CORTADO PARA CIERTAS AGENCIAS, UNIDADES DE NEGOCIOS Y MOSTRANDO EL GRAN TOTAL.....	158
FIGURA 3.79 REPORTE DE DETALLE SINIESTROS ENERO 2006 PRIMERA PÁGINA.....	160
FIGURA 3.80 REPORTE DE DETALLE SINIESTROS ENERO 2006 ÚLTIMA PÁGINA.....	161
FIGURA 3.81 INFORME DE WEB INTELLIGENCE PARA CUADRE DE SINIESTROS.....	161
FIGURA 4.1 PORTAL DE ENTREGA DE INFORMACIÓN BUSINESSOBJECTS INFOVIEW.....	167
FIGURA 4.2 PANTALLA DE ACCESO AL REPORTE DE RESULTADO TÉCNICO.....	169
FIGURA 4.3 RESULTADO TÉCNICO PARA TODOS LOS AÑOS (2005 – 2007).....	169
FIGURA 4.4 RESULTADO TÉCNICO EXPLORANDO EL AÑO 2005.....	170
FIGURA 4.5 RESULTADO TÉCNICO EXPLORANDO EL AÑO 2006.....	170
FIGURA 4.6 RESULTADO TÉCNICO EXPLORANDO EL AÑO 2007.....	171
FIGURA 4.7 INFORME DE INGRESOS POR PRIMAS PARA TODOS LOS AÑOS.....	171
FIGURA 4.8 INFORME DE EGRESOS POR SINIESTROS PARA TODOS LOS AÑOS.....	172
FIGURA 4.9 PANTALLA DE ACCESO AL REPORTE DE INDICADORES DE ANÁLISIS GLOBALES.....	172
FIGURA 4.10 INDICADORES DE ANÁLISIS GLOBALES PARA TODOS LOS AÑOS.....	173
FIGURA 4.11 INDICADORES DE ANÁLISIS GLOBALES PARA EL AÑO 2006 Y EL AGENTE DE SEGUROS AGENCIA ASESORA PRODUCTORA DE SEGUROS 17.....	173

FIGURA 4.12 INDICADORES DE ANÁLISIS GLOBALES PARA EL AÑO 2007 Y LA AGENCIA NORTE 1.....	174
FIGURA 4.13 INDICADORES DE ANÁLISIS GLOBALES PARA EL AÑO 2007 Y LA MARCA CHEVROLET	174
FIGURA 4.14 PANTALLA DE ACCESO AL REPORTE DE ANÁLISIS DE TASAS	175
FIGURA 4.15 TASA PROMEDIO POR AÑO Y MES DE FORMA TABULAR	175
FIGURA 4.16 TASA PROMEDIO POR AÑO Y MES CON GRÁFICA DE TENDENCIAS	176
FIGURA 4.17 TASA PROMEDIO POR MARCA DE VEHÍCULO	176
FIGURA 4.18 TASA PROMEDIO POR MARCA DE VEHÍCULO PARA EL AÑO 2007	177
FIGURA 4.19 TASA PROMEDIO PARA LA MARCA CHEVROLET EN EL AÑO 2007 PROFUNDIZANDO EN TODOS SUS MODELOS (DRILL-DOWN)	177
FIGURA 4.20 PANTALLA DE ACCESO AL REPORTE DE ANÁLISIS DE RECLAMOS	178
FIGURA 4.21 NÚMERO DE VEHÍCULOS SINISTRADOS POR AGENCIA Y AÑO.....	178
FIGURA 4.22 VALOR RECLAMO PROMEDIO POR AGENCIA Y AÑO.....	179
FIGURA 4.23 NÚMERO DE VEHÍCULOS SINISTRADOS POR MES Y AÑO DE FORMA TABULAR	179
FIGURA 4.24 NÚMERO DE VEHÍCULOS SINISTRADOS POR MES Y AÑO EN FORMA DE GRÁFICA DE TENDENCIAS	180
FIGURA 4.25 NÚMERO DE VEHÍCULOS SINISTRADOS POR CAUSA Y AÑO DE FORMA TABULAR	180
FIGURA 4.26 PORCENTAJE DEL NÚMERO DE VEHÍCULOS SINISTRADOS POR CAUSA PARA TODOS LOS AÑOS (2005 – 2007) EN FORMA DE GRÁFICA DE PASTEL	181
FIGURA 4.27 PANTALLA DE ACCESO AL TABLERO DE CONTROL DE SEGUROS	182
FIGURA 4.28 TABLERO DE CONTROL Y ANALÍTICAS UTILIZADAS	183
FIGURA 4.29 TENDENCIA DE MÉTRICAS INTERACTIVA PARA RESULTADO TÉCNICO POR AGENCIA.....	184
FIGURA 4.30 BARÓMETRO DEL PORCENTAJE DE SINISTRALIDAD POR AGENCIA.....	184
FIGURA 4.31 BARÓMETRO DE TASA PROMEDIO POR AGENCIA.....	185
FIGURA 4.32 BARÓMETRO DEL PORCENTAJE DE COMISIONES PAGADAS POR AGENTE DE SEGUROS	185
FIGURA 4.33 GRÁFICO DE PARETO PARA EL RESULTADO TÉCNICO POR MARCA.....	186

CAPÍTULO I

1 MARCO TEÓRICO

1.1 INTELIGENCIA DE NEGOCIOS

1.1.1 INTRODUCCIÓN A LA INTELIGENCIA DE NEGOCIOS

La inteligencia de negocios, correspondiente a la traducción del término en inglés Business Intelligence y cuya abreviación es BI consiste en el conjunto de actividades que una empresa realiza para comprender su situación actual, en dónde ha estado y predecir que sucederá, tomando decisiones estratégicas, tácticas y operativas en base al análisis de sus datos históricos.

Un ambiente de inteligencia de negocios debe estar enfocado a toda la empresa y estar apoyado por una arquitectura que sea sólida, flexible y escalable para soportar los diferentes tipos de análisis y herramientas de inteligencia de negocios que existen y aparecerán en el futuro. Las diferentes arquitecturas de BI han sido propuestas por los padres del Data Warehousing, Bill Inmon y Ralph Kimball y por otros autores que han consolidado las propuestas de estos dos autores.

Como es conocido el activo más importante de una empresa es sus datos. ¿Por qué no aprovechar éstos datos para generar información y obtener el conocimiento necesario para lograr que la empresa sea más competitiva y eficiente; además, por qué no apoyar a la toma de decisiones de las diferentes líneas del negocio a través del análisis de éstos datos? Este es uno de los objetivos principales de la inteligencia de negocios inteligentemente aplicada a una empresa.

La inteligencia de negocios no es sinónimo de Data Warehousing, ésta va mucho más allá. Aunque un Data Warehouse es la principal fuente de datos de un

sistema de inteligencia de negocios, la inteligencia de negocios no necesariamente se limita a ese concepto, tiene un alcance mayor. Las principales actividades de la inteligencia de negocios se agrupan en el reporte, el análisis OLAP (Online Analytical Processing) y el Data Mining.

Tanto la inteligencia de negocios como sus herramientas han ido evolucionando y se han convertido en verdaderas plataformas que soportan las diferentes arquitecturas de inteligencia de negocios. Estas herramientas contienen componentes para la integración de datos, reporte y entrega de información, la misma que puede ser visualizada a través de teléfonos móviles, navegadores de Internet, correo electrónico y aplicaciones de escritorio.

La inteligencia de negocios sigue evolucionando y poco a poco está formando parte de la estructura organizativa de una empresa, como es el caso de los centros de competencias de inteligencia de negocios (BICC, Business Intelligence Competency Center) y a nuevos conceptos abstractos como es el Business Intelligence 2.0 que se refiere al conjunto de herramientas y tecnologías actuales de inteligencia de negocios que han evolucionado y han pasado de ser herramientas reactivas a herramientas proactivas [9].

1.1.2 DEFINICIÓN DE INTELIGENCIA DE NEGOCIOS

La inteligencia de negocios (BI) es una colección de actividades que sirven para entender las situaciones del negocio a través de la realización de varios tipos de análisis sobre los datos de una empresa. La inteligencia de negocios apoya a la toma de decisiones del negocio: estratégicas, tácticas y operativas. Además, ayuda a las empresas a tomar las acciones necesarias para mejorar el rendimiento del negocio [5]. Algunas aplicaciones de la inteligencia de negocios son:

- Gestión del rendimiento del negocio (BPM, Business Performance Management): la gestión del rendimiento del negocio a través de la creación de indicadores claves de rendimiento (KPI, Key Performance Indicators)

[5] RAINARDI, Vincent. **Building a Data Warehouse: With Examples in SQL Server.**

[9] WIKIPEDIA, The Free Encyclopedia. **Business Intelligence 2.0.**

permite a las empresas tomar decisiones tácticas y mejorar el rendimiento operacional.

- **Análisis de clientes:** apoyado con los datos de un CRM (Customer Relationship Management) permite hallar cuáles clientes son rentables para la empresa y cuáles no.
- **Análisis estadístico:** en base a muestras de datos permite proyectar ciertas situaciones basadas en la probabilidad de que un evento ocurra o no. Por ejemplo, en una empresa aseguradora se podría realizar un análisis estadístico o actuarial para definir las tasas técnicas que se aplicarán en la emisión de pólizas de vehículos después de haber encontrado a través de un análisis histórico la frecuencia y probabilidad de que un determinado modelo y marca de vehículo sea siniestro o no.
- **Análisis predictivo:** para predecir a través de herramientas y algoritmos de minería de datos (Data Mining) ciertos patrones que identificarían ciertas condiciones futuras que podrían ser riesgosas para un negocio o no. Por ejemplo, en el negocio asegurador, encontrar patrones que identificarían si ciertos grupos de personas que tienen celular son más propensas a accidentarse en su vehículo que aquellas que no lo tienen.

De acuerdo al nivel y a la complejidad de análisis, las actividades de inteligencia de negocios pueden ser agrupadas en tres categorías:

- Reporteo,
- Procesamiento analítico en línea (OLAP), y
- Minería de datos (Data Mining) [5].

1.1.2.1 Reporteo

El reporte tiene como función la entrega de información a los usuarios del negocio a través de varias formas que varían desde reportes hasta tableros de mando integral (dashboards) con sus respectivos indicadores claves de rendimiento (KPIs).

[5] RAINARDI, Vincent. **Building a Data Warehouse: With Examples in SQL Server.**

En el contexto de la inteligencia de negocios, un reporte es un programa que recupera datos de un Data Warehouse, Data Mart u otro almacén de datos y los presenta al usuario a través de cualquier medio, sea éste un navegador de Internet (browser), correo electrónico, dispositivo móvil o aplicación de escritorio para su visualización y análisis. Los reportes pueden estar calendarizados o generarse automáticamente al ocurrir ciertos eventos del negocio preestablecidos, siempre y cuando la herramienta de inteligencia de negocios utilizada soporte ésta funcionalidad.

Los reportes son contruidos de acuerdo a las especificaciones funcionales de los usuarios del negocio, muestran datos importantes de las diversas fuentes de datos del ambiente de inteligencia de negocios para que sean analizados y se puedan comprender ciertas situaciones del negocio. La forma más común de presentación de reportes es la forma tabular que contiene columnas simples y sus datos, otra forma son las tablas cruzadas o matrices. Los reportes apoyan a la toma de decisiones operativas y tácticas.

Los tableros de mando integral (dashboards) presentan indicadores claves de rendimiento (KPIs) en forma de velocímetros, barómetros, tendencias, semáforos, u otra representación visual. Los tableros de mando integral apoyan a la toma de decisiones estratégicas, ya que con un solo vistazo se puede determinar sí existen problemas de rendimiento del negocio o no analizando los diferentes indicadores claves de rendimiento (KPIs).

1.1.2.2 OLAP (Online Analytical Processing)

Es la actividad que consiste en el análisis interactivo de las transacciones del negocio almacenadas en un Data Warehouse, Data Mart u otro almacén de datos para tomar decisiones del negocio tácticas y estratégicas.

El análisis OLAP es usualmente realizado por los analistas del negocio, administradores del negocio y ejecutivos. Algunas funciones OLAP incluyen las agregaciones resumidas (ROLL-UP), agregaciones detalladas (DRILL-DOWN), el corte y la rotación de los cubos OLAP (SLICE AND DICE), entre las más importantes.

La funcionalidad OLAP dentro de un ambiente de inteligencia de negocios puede ser entregada usando una base de datos relacional o multidimensional. Sí se utiliza una base de datos relacional para el análisis OLAP, éste tipo de análisis se lo conoce como ROLAP (Relational Online Analytical Processing). Sí se utiliza una base de datos multidimensional para el análisis OLAP, éste tipo de análisis se lo conoce como MOLAP (Multidimensional Online Analytical Processing). Finalmente, sí se conjuga el análisis ROLAP y MOLAP, se logra un tipo análisis híbrido conocido como HOLAP (Hybrid Online Analytical Processing).

1.1.2.3 Minería de datos (Data Mining)

La minería de datos o Data Mining es un proceso que consiste en la exploración y búsqueda de patrones y relaciones que describen los datos prediciendo sus valores futuros o inciertos a través de la comprensión de las situaciones del negocio pasadas o históricas. La minería de datos puede ser realizada a través de ciertas técnicas como la caracterización de datos, discriminación de datos, análisis de asociaciones, clasificación de datos, agrupación de datos (clustering), predicciones, análisis de tendencias, análisis de desviaciones y análisis de similitudes.

Cuando la minería de datos es usada para explicar la situación actual o pasada del negocio se la conoce como análisis descriptivo. Al contrario, cuando la minería de datos es utilizada para predecir el futuro del negocio se la conoce como análisis predictivo.

Dentro del ambiente de inteligencia de negocios, algunas aplicaciones populares de la minería de datos consisten en la detección de fraudes, predicción de presupuestos, análisis de clientes potenciales, análisis de tiempos de entrega y en el negocio asegurador, el análisis de perfiles de clientes riesgosos (Customer Risk Profiling).

1.1.3 ARQUITECTURAS DE INTELIGENCIA DE NEGOCIOS

Para lograr un ambiente de inteligencia de negocios dentro de una empresa, es necesario definir e implantar una arquitectura de inteligencia de negocios que

permita a los usuarios de los diferentes niveles organizacionales tomar decisiones basadas en el análisis de los datos que la arquitectura provea.

Los padres del Data Warehousing, Bill Inmon y Ralph Kimball, proponen una arquitectura de inteligencia de negocios similar pero con diferentes variantes para la implantación de proyectos de inteligencia de negocios en una empresa. Por un lado, Inmon propone una arquitectura de inteligencia de negocios llamada CIF (Corporate Information Factory) o fábrica de información corporativa. Por otro lado, Kimball propone una arquitectura de inteligencia de negocios conocida como MD (MultiDimensional) o multidimensional. Las dos propuestas de arquitectura de inteligencia de negocios tienen el mismo propósito pero su principal diferencia radica en la forma de modelar los datos. Inmon sugiere utilizar el modelo entidad-relación (E-R) para el diseño del Data Warehouse y el modelo de datos dimensional para el diseño de los Data Marts; Kimball sugiere utilizar el modelo de datos dimensional para el diseño del Data Warehouse y los Data Marts. [4][5]

1.1.3.1 Corporate Information Factory (CIF)

La fábrica de información corporativa (CIF) es una arquitectura conceptual que describe y categoriza diferentes almacenes de datos enfocados en los procesos organizacionales con un enfoque TOP - DOWN. Estos almacenes de datos son usados para operar y administrar una infraestructura de inteligencia de negocios robusta y eficiente. Los almacenes de datos se enfocan en tres niveles de procesos organizacionales:

- Procesos de operación del negocio: corresponden a las operaciones del día a día del negocio. Estas operaciones están compuestas por datos externos y sistemas OLTP (Online Transactional Processing), sus procesos son estáticos y constituyen la espina dorsal (backbone) de la empresa.
- Procesos de inteligencia de negocios: corresponden a los procesos de búsqueda para comprender mejor a la empresa, sus productos, sus clientes y las diferentes variables del negocio. Aunque los procesos del negocio son estáticos, los procesos de inteligencia de negocios evolucionan

[4] IMHOFF, Claudia; GALEMMO, Nicholas; GEIGER, Jonathan G. **Mastering Data Warehouse Design Relational and Dimensional Techniques.**

[5] RAINARDI, Vincent. **Building a Data Warehouse: With Examples in SQL Server.**

constantemente. Los procesos de inteligencia de negocios apoyan a la toma de decisiones estratégicas.

- Procesos de gestión del negocio: el conocimiento que es generado por los procesos de inteligencia de negocios es institucionalizado e introducido en las operaciones diarias del negocio para toda la empresa. Los procesos de gestión del negocio apoyan a la toma de decisiones tácticas.

La fábrica de información corporativa (CIF) segrega los datos en cinco almacenes de datos que corresponden a los sistemas operacionales, el Data Warehouse, los almacenes de datos operacionales (ODS, Operational Data Store), los Data Marts, y los Marts Operacionales. Cada almacén de datos tiene el rol de proveer la información correcta, al tiempo correcto, en el lugar correcto y en la forma correcta. Además, la fábrica de información corporativa incorpora procesos para mover los datos de los sistemas fuente a los usuarios del negocio de forma eficiente. En la Figura 1.1 se muestra la fábrica de información corporativa (CIF), propuesta por Bill Inmon.

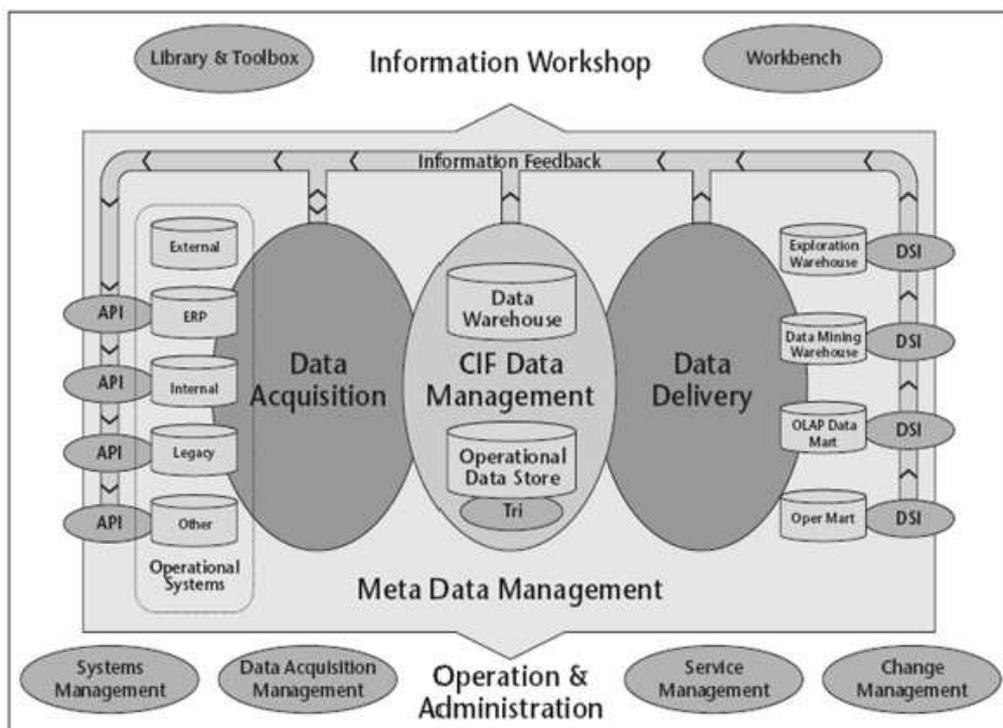


Figura 1.1 La fábrica de información corporativa (CIF) de Bill Inmon [4]

La fábrica de información corporativa (CIF) realiza dos grandes funciones: la obtención de datos de entrada y la entrega de información de salida, cada función tiene sus respectivos componentes y procesos.

- Obtención de datos de entrada: consiste en los procesos y bases de datos involucradas en la adquisición de datos de los sistemas operacionales para integrarlos, limpiarlos y cargarlos en una base de datos para su utilización. Los componentes que se encuentra en ésta función son:

Los sistemas fuente y bases de datos operacionales que contienen los datos de las operaciones diarias del negocio. Este componente es estático por naturaleza y representa la primera aplicación de las reglas del negocio.

El Data Warehouse que es un repositorio de datos integrado, detallado e histórico que apoya a la toma de decisiones estratégicas. El Data Warehouse entrega una vista común de los datos empresariales.

El almacén de datos operacional (ODS) que es un repositorio de datos integrado, detallado y actual que apoya a la toma de decisiones tácticas. Los datos del almacén de datos operacional son volátiles o actualizables y son tan recientes como la tecnología lo permita.

La adquisición de datos es un conjunto de procesos y programas que extraen datos de los sistemas operacionales, los limpian, integran y transforman en un formato empresarial para depositarlos en el Data Warehouse y los almacenes de datos operacionales. Este es el componente más complejo de la fábrica de información corporativa. Muchos de los procesos de adquisición de datos son apoyados por herramientas de extracción, transformación y carga (ETL).

- Entrega de información de salida: consiste en los procesos y bases de datos involucradas en la entrega de inteligencia de negocios a los usuarios y analistas del negocio. Los componentes que se encuentran en ésta función son:

Los Data Marts que son derivados del Data Warehouse y son utilizados para realizar análisis estratégico. Aquí es dónde se realizan la mayor cantidad de actividades analíticas.

Los Marts Operacionales que son derivados de los almacenes de datos operacionales (ODS) y son utilizados para proveer acceso dimensional a los datos operacionales actuales.

La entrega de datos es un conjunto de procesos que mueven los datos del Data Warehouse a los diferentes Data Marts y Marts operacionales. La entrega de datos tiene que ser flexible para soportar todas las herramientas de inteligencia de negocios.

La fábrica de información corporativa (CIF) además de las funciones de obtención de datos y entrega de información realiza las siguientes funciones:

- Gestión de los metadatos: es el conjunto de procesos para coleccionar, administrar e implementar metadatos a través de toda la fábrica de información corporativa. Se agrupan tres categorías de metadatos, los metadatos técnicos, del negocio y administrativos.
- Retroalimentación de información: es un mecanismo para compartir la inteligencia y conocimiento recolectado por la fábrica de información corporativa a través de todos los almacenes de datos. Tiene como objetivo el crear una “Organización del Aprendizaje.”
- Talleres (Workshop) de información: es el conjunto de herramientas disponibles para los usuarios del negocio que permiten la utilización de los recursos de la fábrica de información corporativa. Puede ser un portal Web de una intranet u otras aplicaciones de software.
- Operaciones y administración de la fábrica de información corporativa: se gestionan los cambios, el sistema de base de datos, los metadatos, en general toda la fábrica de información corporativa.

1.1.3.2 Arquitectura Multidimensional (MD)

La arquitectura multidimensional (MD) tiene la premisa de que todos los tipos de análisis de inteligencia de negocios tienen su fundamento en un diseño de datos multidimensional. El esquema estrella constituye el diseño más popular, elegante y útil para la toma de decisiones estratégicas. Una de las principales diferencias entre las dos arquitecturas CIF y MD consiste en el modelamiento de datos. Otro aspecto importante de ésta arquitectura es que su enfoque es BOTTOM – UP.

La arquitectura multidimensional (MD) está dividida en dos grupos de componentes y procesos. Estos grupos de componentes son el Back Room y el Front Room. En el Back Room se realizan los procesos de adquisición y almacenamiento de datos, este componente es equivalente a la función de obtención de datos de entrada de la fábrica de información corporativa (CIF) con algunas diferencias. En la Figura 1.2 se muestra la arquitectura multidimensional (MD):

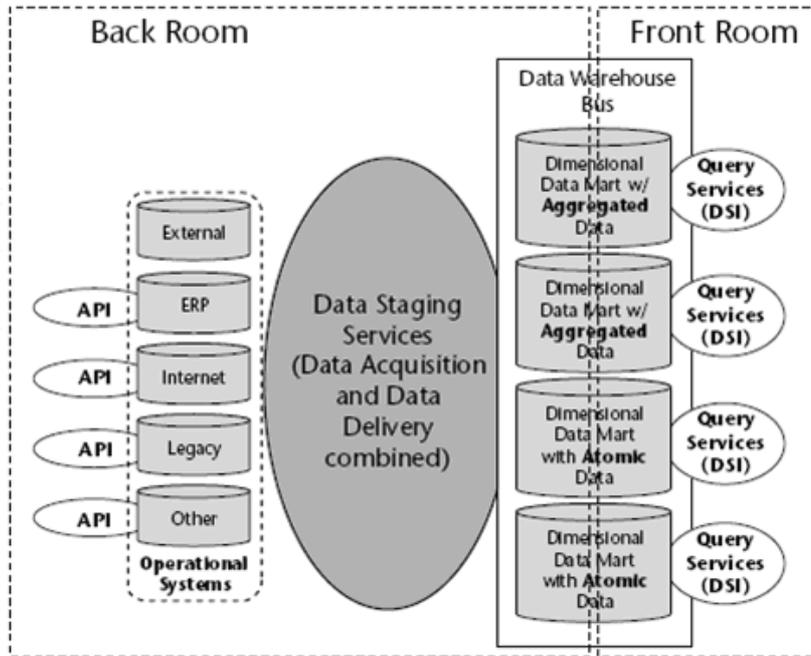


Figura 1.2 La arquitectura multidimensional (MD) de Ralph Kimball [4]

En el Back Room se encuentran los sistemas operacionales y los datos de éstos son transportados a un área de almacenamiento de datos conocida como Stage Area. Esta área contiene las dimensiones conformadas y es el lugar en donde las claves subrogadas (SK, Subrogate Keys) son generadas, aquí también es donde ocurren la mayor parte del procesamiento de las herramientas ETL.

En la arquitectura multidimensional (MD) no es necesario crear un Data Warehouse sino que este se compone de la integración de todos los Data Marts

[4] IMHOFF, Claudia; GALEMMO, Nicholas; GEIGER, Jonathan G. **Mastering Data Warehouse Design Relational and Dimensional Techniques.**

atómicos o detallados y Data Marts agregados, este esquema se lo conoce como la arquitectura de Data Warehouse en bus (DWBA, Data Warehouse Bus Architecture).

La arquitectura multidimensional (MD) comienza y termina enfocada primeramente en unidades de negocio individuales con un requerimiento específico de inteligencia de negocios. Además, la necesidad de crear Data Marts atómicos o Data Marts agregados depende de los requerimientos del negocio y del rendimiento esperado.

El Front Room es la interface para la comunidad del negocio y tiene como objetivo ocultar la complejidad del componente Back Room. Este componente es equivalente a la función de entrega de información de salida en la fábrica de información corporativa (CIF). Los Data Marts atómicos y agregados son la fuente de datos para las herramientas de análisis que consisten de motores OLAP, herramientas de consulta, reporte y minería de datos.

La arquitectura multidimensional (MD) posee también de operaciones de gestión y administración de los metadatos, las fuentes de datos y los grupos de componentes en general.

1.1.3.3 Comparativa entre CIF y la arquitectura multidimensional (MD)

Tanto la arquitectura multidimensional (MD) como la fábrica de información corporativa (CIF) tienen las funciones de obtención de datos y entrega de información. La principal diferencia radica en el modelamiento de datos. La arquitectura CIF propone modelar el Data Warehouse usando diagramas entidad-relación mientras que la arquitectura MD propone usar un modelamiento de datos dimensional para modelar los diferentes Data Marts.

La arquitectura CIF tiene un enfoque TOP – DOWN, es decir tiene como primer objetivo proveer una vista empresarial de los datos a través de un Data Warehouse y luego ir hacia abajo construyendo los diferentes Data Marts de las unidades de negocio que se requieran. Al contrario, la arquitectura MD tiene un enfoque BOTTOM – UP, es decir se parte desde los requerimientos funcionales de cada unidad de negocio o departamento de una empresa construyendo un

Data Mart y así se construyen los Data Marts necesarios para cada departamento. El conjunto de todos los Data Marts constituyen la arquitectura en bus del Data Warehouse (DWBA).

Tanto la arquitectura CIF como la arquitectura MD proveen de datos transaccionales detallados y resumizados. La arquitectura CIF no solo se limita a proveer datos históricos a través de un Data Warehouse o Data Mart de un departamento sino que provee de un almacén de datos operacional (ODS) y Mart operacionales para proveer datos lo más actualizados posibles así como la tecnología lo permita. En la arquitectura multidimensional (MD) también es posible crear almacenes de datos operacionales (ODS) modificando el modelo conceptual obteniendo los datos del Stage Area y cargándolos en el ODS.

En la Figura 1.3 se muestra gráficamente la comparativa entre las dos arquitecturas de inteligencia de negocios CIF y MD, se incluyen algunas comparaciones descritas anteriormente y otras que no son muy relevantes.

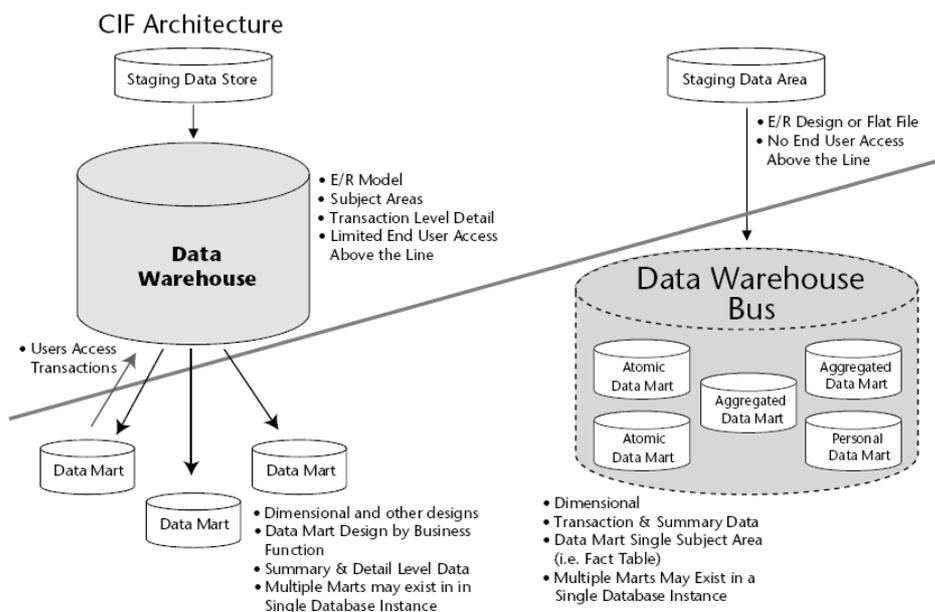


Figura 1.3 Comparativa entre CIF y MD [4]

[4] IMHOFF, Claudia; GALEMMO, Nicholas; GEIGER, Jonathan G. **Mastering Data Warehouse Design Relational and Dimensional Techniques.**

En la Tabla 1.1 se resumen las características más relevantes de comparación entre la arquitectura multidimensional (MD) y la fábrica de información corporativa (CIF).

Comentario [MM1]: Es necesario mencionar porque se seleccionó la metodología de Kimball para el desarrollo del proyecto de titulación.

Costo, tiempo, una sola área del negocio y otras justificaciones.

Característica	Fábrica de Información Corporativa (CIF)	Arquitectura Multidimensional (MD)
Técnica de modelamiento de datos	Relacional y Dimensional	Dimensional
Enfoque	TOP – DOWN	BOTTOM – UP
Es necesario la creación de un Data Warehouse	Sí	No
Los Data Mart son derivados del Data Warehouse	Sí	No
Provee una vista empresarial de los datos	Sí a través de un Data Warehouse	Parcialmente a través de la arquitectura de Data Warehouse en bus
Provee datos detallados	Sí	Sí
Provee datos resumizados	Sí	Sí
Es libre de procesos (process-free)	Sí	No
Tiempo de Implantación	Alto	Bajo
Costo de implantación	Alto	Bajo
Soporta almacenes de datos operacionales (ODS) y Marts Operacionales	Sí	No conceptualmente, pero prácticamente se los puede implementar

Tabla 1.1 Tabla comparativa entre CIF y MD

1.1.4 HERRAMIENTAS DE INTELIGENCIA DE NEGOCIOS

Para poder implantar una arquitectura de inteligencia de negocios en una empresa de forma exitosa es necesario utilizar una herramienta o conjunto de herramientas que apoyen en todos los procesos de la arquitectura que van desde la obtención de datos hasta la entrega de información a los usuarios del negocio.

Existen diversos tipos de herramientas de software para inteligencia de negocios orientadas a una función específica como por ejemplo las herramientas de extracción, transformación o carga (ETL), herramientas de reporte, herramientas de análisis OLAP, herramientas de minería de datos y herramientas para visualizar tableros de mando integral (dashboards). Muchas de las herramientas de inteligencia de negocios se han convertido en plataformas para apoyar a todas las funciones de una arquitectura de inteligencia de negocios, constituyen una suite de herramientas que apoyan en los procesos ETL, gestión de metadatos, diseño de modelos dimensionales, reporte, tableros y minería de datos.

La mayoría de herramientas de inteligencia de negocios son comerciales y constituyen verdaderas plataformas para apoyar a un ambiente de inteligencia de negocios en una empresa. Algunas de las herramientas comerciales más conocidas de inteligencia de negocios son: BusinessObjects [10] , Cognos [11] , Oracle Hyperion [12] , SAS [13] y Microsoft Analysis Services [14] . También existen herramientas de código abierto (Open Source) muy conocidas como es el caso de Pentaho [15] y Eclipse BIRT (Business Intelligence and Reporting Tools) [16].

La plataforma de inteligencia de negocios que se usará en el proyecto de titulación es la de BusinessObjects cuyo nombre comercial es BusinessObjects Enterprise XI Release 2. En el siguiente subcapítulo se explica de forma muy

[10] BusinessObjects, an SAP Company. BusinessObjects XI 3.0.

[11] Business Intelligence and Performance Management products from Cognos, an IBM Company.

[12] Oracle and Hyperion.

[13] SAS, Business Intelligence Software and Predictive Analysis.

[14] Microsoft SQL Server: Analysis Services.

[15] Pentaho Comercial Open Source Business Intelligence.

[16] Eclipse BIRT Home.

general la arquitectura de la herramienta y cómo apoyará en las diferentes fases de proyecto.

1.1.4.1 BusinessObjects Enterprise XI Release 2

BusinessObjects Enterprise XI Release 2 (BOE XI R2) es una plataforma de inteligencia de negocios probada y líder en la industria de herramientas de inteligencia de negocios. BOE XI R2 es una plataforma de alto rendimiento y confiabilidad que provee escalabilidad debido a que tiene una arquitectura orientada a servicios (SOA, Service Oriented Architecture), seguridad ya que soporta protocolos de red seguros, y confiabilidad ya que es tolerante a fallos. [7]

BOE XI R2 es una suite de productos integrados de inteligencia de negocios que apoyan a realizar un seguimiento del rendimiento del negocio, comprender el negocio, gestionar el negocio y lograr la estandarización de la inteligencia de negocios al usar toda la suite de herramientas de BI. Las herramientas de la suite BOE XI R2 sirven para reporte, análisis de consultas, gestión del rendimiento, e integración de datos.

En la Figura 1.4 se muestra la plataforma BOE XI R2 y sus diversos componentes que tienen como fin el apoyar a la estandarización de la inteligencia de negocios y tener una arquitectura de inteligencia de negocio escalable, segura y confiable.



Figura 1.4 La plataforma de BI Business Objects Enterprise XI Release 2 [7]

[7] CLARK, P. Business Objects Products and Services: the Industry's Leading Suite of Integrated Business Intelligence Products.

Como se puede observar en la Figura 1.4 la plataforma BOE XI R2 contiene todos los componentes necesarios para implementar cualquiera de las arquitecturas de inteligencia de negocios (CIF o MD) descritas anteriormente. El primer componente es el de integración de datos (Data Integration), el cuál permite realizar la función de obtención de datos de entrada. El siguiente componente es la plataforma de inteligencia de negocios, la cuál gestiona los metadatos, provee de un portal de entrega de información de BI, permite la colaboración, contiene motores de analíticas, provee de seguridad y los servicios necesarios para que la arquitectura de BI funcione correctamente. Los siguientes componentes son el de reporte y análisis de consultas, en dónde es posible crear reportes empresariales, realizar análisis interactivo y consultas Ad Hoc. El último componente de BOE XI R2 tiene como función la gestión del rendimiento proveyendo a los usuarios del negocio de tableros de mando integral, scorecards para una implementación de la técnica gerencial Balanced Scorecard [17], y aplicaciones de gestión del rendimiento.

1.1.5 TENDENCIAS EN LA INTELIGENCIA DE NEGOCIOS

Las herramientas de inteligencia de negocios y el estado del arte de la inteligencia de negocios (BI) han evolucionado bastante durante los últimos 20 años y continúan evolucionando. Han aparecido nuevas herramientas y tecnologías consolidadas como plataformas de inteligencia de negocios, tal es el caso de BOE XI R2, y siguen apareciendo nuevas herramientas y plataformas proactivas, las cuáles usan tecnologías de aplicaciones ricas de Internet (Web 2.0) para la entrega de información a los usuarios del negocio.

Nuevos conceptos y definiciones han ido apareciendo en el estado del arte de la inteligencia de negocios, tal es el caso de los centros de competencias de inteligencia de negocios (BICC, Business Intelligence Competency Center) que buscan crear departamentos o unidades organizacionales enfocadas al apoyo de las iniciativas de la inteligencia de negocios y su utilización dentro de una empresa. Muchas empresas están yendo hacia allá de tal forma que la inteligencia de negocios en una empresa sea consolidada como un departamento

[17] The Balanced Scorecard Institute.

o unidad de negocio y se convierta en el eje fundamental que apoye a la toma de decisiones estratégicas, tácticas y operacionales.

1.1.5.1 BICC, Business Intelligence Competency Center

Según Claudia Imhoff, presidenta de Intelligent Solutions Inc. [18], “La mayoría de organizaciones se acercan a una iniciativa de inteligencia de negocios como un simple proyecto desconectado. El acercamiento tradicional ha sido el de reunir un grupo de desarrolladores, construir una solución para un proyecto de BI y desagrupar el equipo cuando el proyecto termina. El desconectar los proyectos e iniciativas de inteligencia de negocios ocasiona que las compañías fallen en crear un ambiente de BI confiable y consistente.” En resumen no hay más de inteligencia de negocios que simplemente implementar tecnología careciendo de un acercamiento comprensivo y estratégico para diseñar, implementar, administrar, controlar y monitorear iniciativas de inteligencia de negocios.

Para superar muchos de los obstáculos como los que menciona Imhoff muchas empresas han creado los centros de competencias de inteligencia de negocios (BICC). Un BICC es un equipo multifuncional que posee una estructura organizacional formal, permanente y que cumple con las siguientes actividades:

- Planear y priorizar iniciativas de inteligencia de negocios,
- Gestionar y dar soporte a las iniciativas de inteligencia de negocios, y
- Promover el uso amplio de la inteligencia de negocios a través de toda la organización.

Un BICC contiene representantes especialistas en las tecnologías de la información y representantes del negocio auspiciados por los ejecutivos de alto nivel. Las personas que trabajan en un BICC tienen que asegurar que las estrategias corporativas se alineen con las estrategias de la inteligencia de negocios.

Un BICC puede ser una compañía externa, otro departamento, o una unidad de negocio virtual con representantes que estén encargados de otras actividades.

[18] Intelligent Solutions, Inc.

Cualquiera que sea la forma que adopte un BICC, ésta será un ejemplo positivo de cambio cumpliendo con los siguientes roles:

- Líderes de iniciativas de inteligencia de negocios y su alineación con los objetivos empresariales.
- Expertos avanzados de analíticas y minería de datos.
- Expertos en adquisición e integración de datos.
- Regularizadores de datos.
- Desarrolladores de aplicaciones de inteligencia de negocios.
- Entrenadores, capacitadores y especialistas en soporte técnico.

Algunos de los beneficios que un BICC provee a una empresa son:

- Aumento del uso de la inteligencia de negocios para apoyar objetivos estratégicos.
- Incremento de la satisfacción de los usuarios del negocio.
- Mejor comprensión del valor de la inteligencia de negocios para el crecimiento empresarial.
- Incremento en la velocidad de toma de decisiones.
- Bajos costos del personal y de software.

Además de estos beneficios, un BICC tiene como función facilitar la transferencia del conocimiento y mejorar las habilidades analíticas de una empresa, permitiéndole:

- Preservar y explotar el valor completo de las inversiones en tecnología,
- Establecer procesos y estándares de BI y analíticas del negocio,
- Reaccionar rápidamente y proactivamente a los cambios del negocio,
- Reducir el riesgo en desarrollar nuevos productos o mercados, y
- Asegurarse que el conocimiento generado por la inteligencia de negocios sea compartido en toda la empresa. [2]

[2] DAVIS, Jim; MILLER, Gloria; RUSSELL, Allan. **Information Revolution Using the Information Evolution Model to Grow Your Business.**

1.2 DATA WAREHOUSING

1.2.1 PERSPECTIVA HISTÓRICA

En los años 70 las primeras aplicaciones comerciales fueron construidas para computarizar las operaciones del día a día de un negocio, estas aplicaciones funcionaban sobre mainframes y eran extremadamente costosas. En aquella época era necesario usar un medio de almacenamiento de acceso rápido por lo que se inventó el almacenamiento en disco, aquí los datos podían ser leídos directamente y no secuencialmente. Esto dio origen a los primeros sistemas de gestión de base de datos (DBMS, Database Management System) jerárquicos y en red, los cuales eran muy complejos de administrar y usar. La facilidad de acceso directo a los datos permitió el apareamiento de los primeros sistemas de procesamiento transaccional en línea (OLTP, Online Transactional Processing).

En los años 80 los datos solían ser almacenados usando el modelo CODASYL (Conference on Data Systems Languages) [19] el cuál era muy complejo y difícil de gestionar. Esta situación cambió con la introducción de las bases de datos relacionales cuyo padre es el Dr. Edgar. F. Codd (1923 – 2003). Oracle fue el primer sistema de gestión de base de datos comercial, el cual fue lanzado al mercado en el año de 1979 [21] con el nombre comercial de Oracle versión 2.

Con la introducción de los computadores personales (PCs) la computación se movió de los mainframes a los sistemas cliente / servidor. Los sistemas OLTP adoptaron ésta arquitectura con la función de coleccionar datos eficientemente, pero el desafío en aquella época era el cómo interpretar los datos coleccionados con el objetivo de realizar análisis sobre éstos. El término de Data Warehouse o bodega de datos fue acuñado por Bill Inmon [22], quien es reconocido ampliamente en el campo como el padre del Data Warehousing. Los Data Warehouses fueron construidos para ver el negocio en el tiempo, determinar tendencias y apoyar a la toma de decisiones administrativas.

[19] WIKIPEDIA, The Free Encyclopedia. **CODASYL**.

[21] WIKIPEDIA, The Free Encyclopedia. **Oracle Corporation**.

[22] WIKIPEDIA, The Free Encyclopedia. **Bill Inmon**.

En un Data Warehouse la actividad primaria que se realiza es la consulta o lectura de datos, la actualización de datos sólo se realiza en los procesos de carga. Un Data Warehouse necesita de un diseño de datos diferente, Bill Inmon propuso que el diseño del Data Warehouse sea realizado usando el modelamiento entidad-relación. Este enfoque fue adoptado al inicio hasta la aparición de Ralph Kimball [23], quien propuso el modelamiento de datos dimensional donde los datos son organizados en tablas de hechos y dimensiones. El modelamiento dimensional ayuda a optimizar el rendimiento de un Data Warehouse y ofrece una representación de datos comprensible para los usuarios del negocio.

Los personajes más influyentes en el campo del Data Warehousing son Bill Inmon y Ralph Kimball, el adoptar la metodología que cada autor propone para la construcción de un Data Warehouse depende de las necesidades del negocio, el tiempo y los recursos disponibles. [3]

1.2.2 DEFINICIÓN DE DATA WAREHOUSE

Según Bill Inmon, “un Data Warehouse es una colección de datos, orientada a un tema, integrada, no volátil y variable en el tiempo que soporta las decisiones administrativas.”

Según Ralph Kimball, “un Data Warehouse es un sistema que extrae, limpia, integra y entrega datos fuente dentro de un almacén de datos dimensional (DDS, Dimensional Data Store) y soporta e implementa consultas de análisis para la toma de decisiones.”

Considerando al Data Warehouse como un sistema que será fuente de datos en una arquitectura de inteligencia de negocios, la definición de Rainardi [5] sobre un Data Warehouse es la que más se alinea con éste propósito y la que conjuga tanto la definición de Inmon como la de Kimball.

[3] HOBBS, Lilian; HILLSON, Susan; LAWANDE, Shilpa; SMITH, Pete. **Oracle Database 10g Data Warehousing**.

[5] RAINARDI, Vincent. **Building a Data Warehouse: With Examples in SQL Server**.

[23] WIKIPEDIA, The Free Encyclopedia. **Ralph Kimball**.

Según Vincent Rainardi, “un Data Warehouse es un sistema que recupera y consolida datos periódicamente de los sistemas fuente a un almacén de datos dimensional (DDS) o normalizado. Usualmente guarda años de historia y es consultado para inteligencia de negocios u otras actividades analíticas.”

En la Figura 1.5 se ilustra la definición de Rainardi.

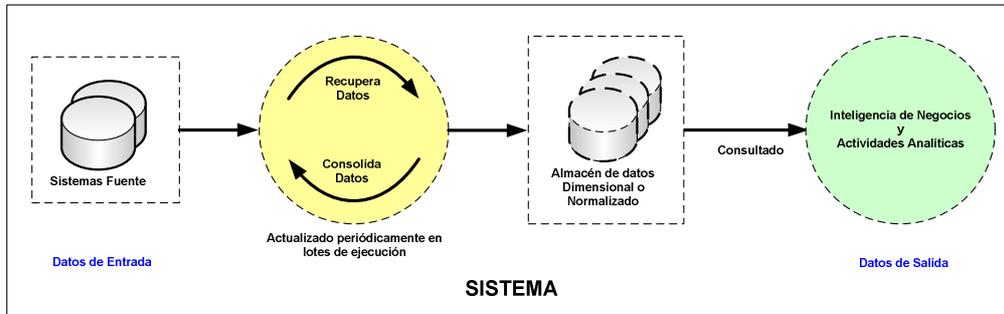


Figura 1.5 Ilustración de la definición de Data Warehouse

En los siguientes subcapítulos se explica cada función del Data Warehouse considerado como un sistema. [5]

1.2.2.1 Recuperación de datos

La recuperación de datos consiste en un conjunto de rutinas conocidas como extracción, transformación y carga (ETL, Extract, Transform and Load). Las rutinas de extracción se encargan de obtener los datos de los sistemas fuente. Las rutinas de transformación sirven para estandarizar datos, derivarlos, validarlos y limpiarlos con el objetivo de tener los datos en un formato común. Las rutinas de carga se encargan de depositar los datos transformados en los almacenes de datos normalizados o dimensionales.

1.2.2.2 Consolidación de datos

Un Data Warehouse consolida datos de varios sistemas transaccionales. Los datos de un Data Warehouse están integrados. Sí existen varias fuentes de datos se tiene que considerar lo siguiente para lograr la integración de datos:

[5] RAINARDI, Vincent. **Building a Data Warehouse: With Examples in SQL Server.**

- Disponibilidad de datos. No todos los datos de un sistema pueden existir en otro. Por ejemplo, en el sistema A se tienen los datos de calle secundaria, principal, número y barrio para una dirección. En el sistema B se tiene sólo dirección y barrio. Para lograr integrar los datos de los sistemas A y B, en el sistema A la dirección estaría conformada sólo por dirección y barrio que son los datos que contiene el sistema B, por lo que:

[Sistema A].[Dirección] = [Sistema A].[Calle Principal] + [Sistema A].[Calle Secundaria] + [Sistema A].[Número]

Concatenando los datos de dirección en el sistema A, se pueden integrar los datos de dirección del sistema B que únicamente son dirección y barrio.

- Rangos de tiempo. Es necesario considerar los cálculos que cada sistema fuente realiza en base a fechas diferentes. Por ejemplo, el sistema A podría calcular un concepto del negocio semanalmente, y el sistema B podría calcular el mismo concepto diariamente.
- Diferentes definiciones. Es muy probable que varios sistemas fuente tengan definiciones distintas para un mismo concepto del negocio. Por ejemplo, el sistema A podría representar el subtotal de venta de una factura como la suma del subtotal más los descuentos mientras que el sistema B podría representar el subtotal únicamente como la suma de todos los ítems vendidos en una factura sin los descuentos.
- Conversiones. Se tiene que considerar las diferentes unidades de medida que cada sistema fuente maneja. Por ejemplo, cuando se integran datos de sistemas heredados es necesario considerar conversiones como el cambio de moneda. En nuestro caso sería necesario convertir los datos de sucres a dólares.
- Coincidencia de datos (matching). Es necesario que los datos de varios sistemas fuente coincidan para lograr consistencia de datos. Por ejemplo, un mismo cliente podría aparecer duplicado en el Data Warehouse, si es que no se realizan las operaciones de equivalencia de datos en los sistemas fuente. Los clientes podrían ser identificados en los sistemas fuente usando la combinación de cédula, apellidos y nombres.

1.2.2.3 Datos actualizados periódicamente

En un Data Warehouse es necesario determinar el período de recuperación y consolidación de datos. Este período tiene que estar basado en los requerimientos del negocio y la frecuencia de actualización de datos de los sistemas fuente.

Los períodos de actualización podrían variar desde actualizaciones en línea, diarias, semanales, mensuales o anuales. Es importante considerar también que las herramientas ETL que se utilicen en la arquitectura de inteligencia de negocios permitan manejar éstos períodos de actualización requeridos. BusinessObjects Data Integrator [24] es una herramienta ETL que permite la actualización de datos en línea. Oracle Database 10g posee también una funcionalidad conocida como Asynchronous Change Data Capture [25] que permite conocer los datos de las tablas que recién cambiaron para usarlos en los procesos ETL.

Los datos son actualizados en lotes de ejecución (batches) para preservar el rendimiento de los sistemas fuente.

1.2.2.4 Almacén de datos

Un almacén de datos puede ser dimensional o normalizado. Un almacén de datos dimensional se lo conoce como DDS (Dimensional Data Store), un almacén de datos normalizado se lo conoce como NDS (Normalized Data Store).

Un almacén de datos dimensional (DDS) es una o varias bases de datos que contiene una colección de Data Marts dimensionales. Un DDS está desnormalizado y sus dimensiones están conformadas. Un DDS puede ser implementado físicamente como un esquema estrella (Star Schema), un esquema con copos de nieve (Snowflake) o un esquema constelación (Galaxy Schema o Fact Constellation Schema). Ralph Kimball propone la construcción de varios DDS para lograr una arquitectura de bus de Data Warehouse (DWBA).

[24] BusinessObjects Data Integrator.

[25] Oracle Database 10g Asynchronous Change Data Capture.

Un almacén de datos normalizado (NDS) es una o más bases de datos relacionales con poca o casi nula redundancia de datos. La normalización se logra implementando hasta las cinco formas normales existentes (1FN – 5FN). Un NDS constituye un formato mejor para integrar datos de varios sistemas fuente y es usado principalmente para Data Warehouses empresariales. Bill Inmon propone la construcción de un Data Warehouse como un NDS.

1.2.2.5 Historia

Un Data Warehouse tiene la capacidad de almacenar historia. A diferencia de un sistema OTLP, un Data Warehouse puede almacenar 5 o más años de historia. La cantidad de datos históricos que se almacenarán en un Data Warehouse depende de los requerimientos del negocio.

En un Data Warehouse el almacenamiento de los datos maestros es una de las características principales. Un dato maestro es aquel que describe las transacciones del negocio, responde a las preguntas ¿quién?, ¿qué? y ¿dónde? Además la técnica conocida como dimensión que cambia lentamente (SCD, Slowly Changing Dimension) es usada para preservar la información histórica sobre los datos dimensionales. [8]

1.2.2.6 Consultas

La consulta es un proceso de obtención de datos de un almacén de datos la cuál satisface ciertos criterios. Un Data Warehouse es construido únicamente para ser consultado y no puede ser actualizado, los procesos ETL son los únicos que pueden actualizar un Data Warehouse.

Las consultas son realizadas por las herramientas de inteligencia de negocios con el fin de realizar las diversas actividades analíticas que se requieran, éstas pueden ser el reporte, el análisis OLAP o la minería de datos (Data Mining).

1.2.3 DEFINICIÓN DE DATA MART

Un Data Warehouse contiene múltiples temas que proveen una vista consolidada de la empresa sobre todas las líneas del negocio. Un Data Mart es un Data

[8] KIMBALL, Ralph. *The data warehouse toolkit: the complete guide to dimensional modeling*.

Warehouse de un tema o aplicación específica que contiene datos para una sola línea de negocio como por ejemplo ventas, finanzas o marketing. La mayor diferencia entre un Data Warehouse y un Data Mart es el alcance de la información que contienen.

Debido a que el alcance de un Data Mart es pequeño y los datos son obtenidos de pocas fuentes, el tiempo de implementación es corto a diferencia de la implementación de un Data Warehouse que puede tomar de dieciocho meses a tres años. [3]

Un Data Mart puede ser dependiente o independiente. Un Data Mart dependiente se deriva de un Data Warehouse como lo propone la fábrica de información corporativa de Inmon. Un Data Mart independiente no se deriva de un Data Warehouse, son construidos a partir de los sistemas fuente que pueden ser un sistema OLTP, archivos o sistemas heredados.

En la Figura 1.6 y Figura 1.7 se ilustran los dos tipos de Data Marts.

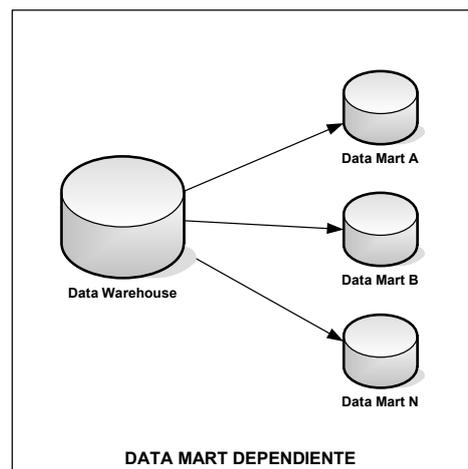


Figura 1.6 Data Mart Dependiente, propuesto por Bill Inmon

[3] HOBBS, Lilian; HILLSON, Susan; LAWANDE, Shilpa; SMITH, Pete. **Oracle Database 10g Data Warehousing**.

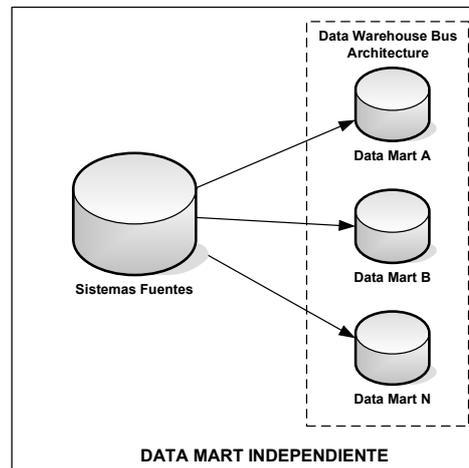


Figura 1.7 Data Mart Independiente, propuesto por Ralph Kimball

La construcción de un Data Mart pospone muchas de las decisiones críticas que después se convierten en necesarias cuando el número de Data Marts crece. Además, evita problemas internos dentro de una organización y utiliza equipos de desarrollo autónomos.

Para resolver los problemas de construcción de un Data Warehouse completo y obtener un retorno de la inversión (ROI, Return on Investment) [26] a tiempo es preferible construir modularmente o incrementalmente el Data Warehouse con Data Marts dependientes siguiendo un alcance por fases. Un alcance por fases provee una retroalimentación constante del negocio, limita el alcance y se planean mejoras cada tres o seis meses.

1.2.4 ARQUITECTURA DE UN DATA WAREHOUSE

Un Data Warehouse conceptualizado como un sistema está conformado por dos arquitecturas principales.

1. Arquitectura de flujo de datos. Esta arquitectura se refiere a cómo los almacenes de datos son organizados dentro del sistema de Data Warehouse y cómo los datos fluyen de los sistemas fuente a los usuarios del negocio.

[26] WIKIPEDIA, The Free Encyclopedia. **Rate of return.**

2. Arquitectura del sistema. Esta arquitectura se refiere a la configuración física de los servidores, redes, software, almacenamiento y computadores clientes. [5]

1.2.4.1 Arquitectura de flujo de datos

Una arquitectura de flujo de datos es una configuración de almacenes de datos dentro de un sistema de Data Warehouse que indica cómo los datos fluyen desde los sistemas fuente a través de los almacenes de datos hasta llegar a las aplicaciones usadas por los usuarios finales.

Un almacén de datos (Data Store) es una o más bases de datos o archivos que contienen datos del Data Warehouse organizados en un formato en particular que están involucrados en los procesos del sistema de Data Warehouse.

Con respecto a su finalidad, los almacenes de datos pueden ser de tres tipos:

- De interface de usuario. Estos almacenes de datos están disponibles para ser consultados por los usuarios y aplicaciones finales.
- Interno. Es un almacén de datos usado internamente por los componentes del sistema de Data Warehouse para los propósitos de integración, limpieza, metadatos, y preparación de datos. Estos almacenes de datos no pueden ser consultados por los usuarios ni aplicaciones finales.
- Híbrido. Es un almacén de datos usado para mecanismos internos del sistema de Data Warehouse y para ser consultado por los usuarios y aplicaciones finales.

Con respecto al formato de los datos, los almacenes de datos pueden ser de cuatro tipos:

- Área de almacenamiento (Stage Area). Es un almacén de datos interno usado para transformar y preparar datos obtenidos de los sistemas fuentes.
- Almacén de datos normalizado (NDS, Normalized Data Store). Es un almacén de datos interno en la forma de una o más bases de datos relacionales y

[5] RAINARDI, Vincent. **Building a Data Warehouse: With Examples in SQL Server.**

normalizadas que integra datos de varios sistemas fuentes y posee datos históricos.

- Almacén de datos operacional (ODS, Operational Data Store). Es un almacén de datos híbrido en la forma de una o más bases de datos relacionales y normalizadas que contiene los datos transaccionales y las versiones más recientes de los datos maestros. No posee datos históricos, sino la versión más reciente de los datos.
- Almacén de datos dimensional (DDS, Dimensional Data Store). Es un almacén de datos de interface de usuario dónde los datos están organizados en un formato dimensional para soportar consultas analíticas.

La arquitectura de flujo de datos es una de las primeras cosas que se tiene que decidir cuando se construye un sistema de Data Warehouse ya que ésta determina los componentes a construir y afecta a los planes y costos del proyecto. Además, el diseño de la arquitectura de flujo de datos debe cumplir con los requerimientos de datos y los requerimientos de calidad del proyecto.

Existen cuatro tipos de arquitectura de flujo de datos:

- Almacén de datos dimensional simple (Simple DDS).
- Almacén de datos normalizado más almacén de datos dimensional (NDS + DDS).
- Almacén de datos operacional más almacén de datos dimensional (ODS + DDS).
- Data Warehouse Federado (FDW, Federated Data Warehouse).

1.2.4.1.1 Arquitectura de flujo de datos general

En la Figura 1.8 se muestra una arquitectura de flujo de datos general para un sistema de Data Warehouse. En esta se incluyen todos los componentes posibles y su descripción. Cabe destacar que las cuatro arquitecturas que se describirán en los siguientes subcapítulos son variaciones de ésta y se relacionan con las propuestas de arquitecturas de inteligencia de negocios de Bill Inmon y Ralph Kimball.

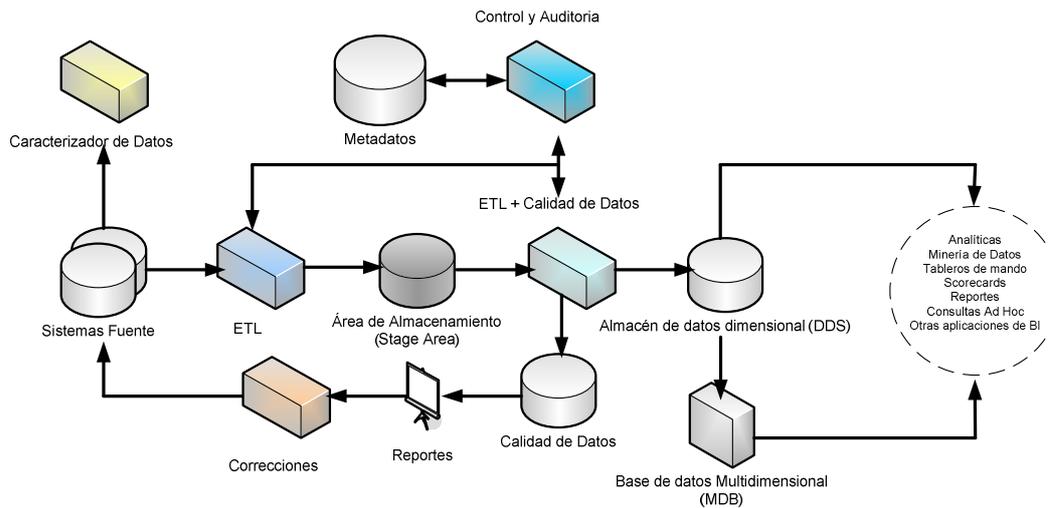


Figura 1.8 Arquitectura de flujo de datos general para un sistema de Data Warehouse [5]

A continuación se describe cada uno de los componentes de la arquitectura de flujo de datos general para un sistema de Data Warehouse:

Los sistemas fuente, son los sistemas OLTP que capturan y almacenan transacciones del negocio. Pueden ser archivos planos u otras fuentes de datos digitalizados.

El caracterizador de datos es una herramienta que tiene la capacidad de analizar los datos encontrando cuantas filas existen en una tabla, cuantos registros tienen columnas con valores nulos y otras formas de caracterizar los datos.

Los procesos ETL se conectan a los sistemas fuente, extraen, transforman y cargan los datos en un sistema destino, que puede ser un área de almacenamiento o un almacén de datos normalizado o dimensional.

El almacén de datos dimensional (DDS) es una base de datos que almacena datos de un Data Warehouse en un formato diferente al del OLTP dónde los datos

[5] RAINARDI, Vincent. **Building a Data Warehouse: With Examples in SQL Server.**

están organizados en un formato dimensional. Un DDS contiene datos integrados de varias fuentes y está optimizado para realizar consultas de análisis.

Las reglas de calidad de datos realizan varios chequeos de calidad de datos. Los datos que no cumplen con éstas reglas son almacenados en una base de datos de calidad de datos para ser corregidos y que puedan ser cargados hasta que cumplan ciertos límites de tolerancia.

La base de datos de metadatos contiene información sobre la estructura de los datos, su significado, su uso, las reglas de calidad y otra información necesaria acerca de los datos.

El componente de auditoría y control registra las operaciones del sistema y su uso dentro de la base de datos de metadatos. Este componente ayuda a comprender los sucesos que ocurrieron durante los procesos ETL.

La base de datos multidimensional (MDB) conocida como cubos, es una forma de base de datos donde los datos se almacenan en celdas y la posición de cada celda está definida por un número de variables llamadas dimensiones. Cada celda del cubo representa un evento del negocio y los valores de las dimensiones muestran ¿cuándo? y ¿dónde? sucedió un evento.

Tanto los almacenes de datos dimensionales (DDS) como las bases de datos multidimensionales (MDB) son consultados por las aplicaciones de inteligencia de negocio que pueden ser reportes, tableros, analíticas, minería de datos, scorecards u otras aplicaciones de inteligencia de negocios.

1.2.4.1.2 Arquitectura simple con un almacén de datos dimensional (Simple DDS)

La arquitectura como su nombre lo dice es simple ya que los datos no tienen que ser cargados en ningún tipo de almacén de datos normalizado. Los datos son extraídos a un área de almacenamiento temporal (Stage Area) y de aquí son extraídos al almacén de datos dimensional (DDS). Es importante destacar que en los procesos ETL para pasar del área de almacenamiento al DDS se aplican reglas de calidad de datos para asegurar su calidad.

En ésta arquitectura es posible crear más almacenes de datos dimensionales (DDS) y es recomendable usarla cuando se tiene una sola fuente de datos y no es necesario integrar datos de otras fuentes. Además, ésta arquitectura se relaciona con la propuesta de Ralph Kimball para lograr una arquitectura de inteligencia de negocios multidimensional (MD) ya que cada DDS tiene un formato de datos dimensional y no es necesario que se deriven de una base de datos normalizada como lo propone Bill Inmon con la fábrica de información corporativa (CIF).

Cabe mencionar también que ésta arquitectura almacena los metadatos tanto de las aplicaciones de inteligencia de negocios como de los procesos ETL a través del componente de control y auditoria. En la Figura 1.9 se muestra la arquitectura simple con un almacén de datos dimensional (DDS).

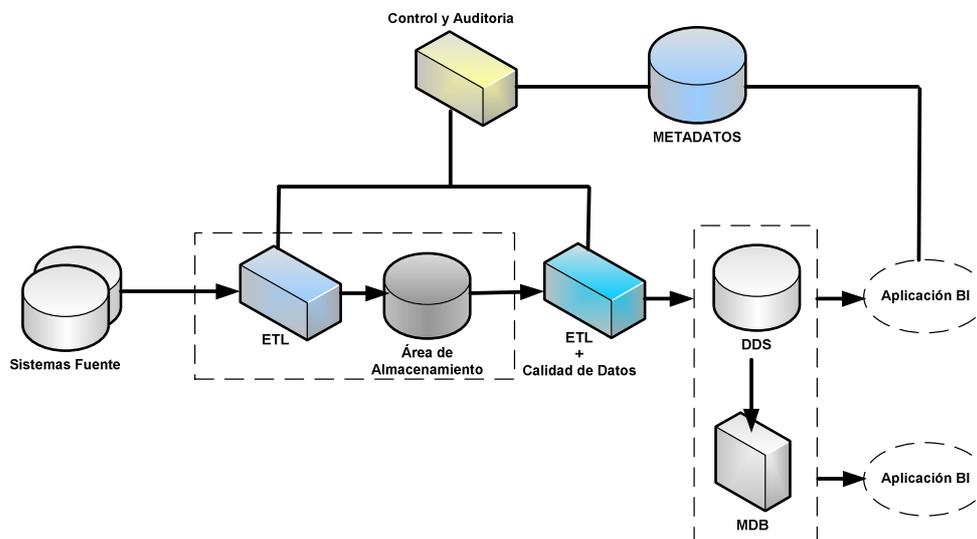


Figura 1.9 Arquitectura simple DDS [5]

1.2.4.1.3 Arquitectura con un almacén de datos normalizado más un almacén de datos dimensional (NDS + DDS)

La arquitectura (NDS + DDS) propone primero la creación de un almacén de datos normalizado (NDS) y la creación de los almacenes de datos dimensionales (DDS) como derivados de éste. Esta arquitectura se relaciona con la propuesta de Bill Inmon para lograr una fábrica de información corporativa (CIF). Como Inmon

[5] RAINARDI, Vincent. **Building a Data Warehouse: With Examples in SQL Server.**

propone, primero se construye un Data Warehouse normalizado y luego se derivan de este los Data Marts que se requieran como almacenes de datos dimensionales.

Cabe destacar que en los procesos ETL para extraer los datos de los sistemas fuente al NDS se aplican reglas de calidad de datos y los datos que no cumplen con las reglas aplicadas son almacenados en un área de almacenamiento temporal (Stage) con el propósito de cargar datos de calidad en el NDS. De igual forma en ésta arquitectura existe el componente de control y auditoria y los metadatos que cumplen la misma función que en todas las arquitecturas de Data Warehouse. En la Figura 1.10 se muestra la arquitectura (NDS + DDS).

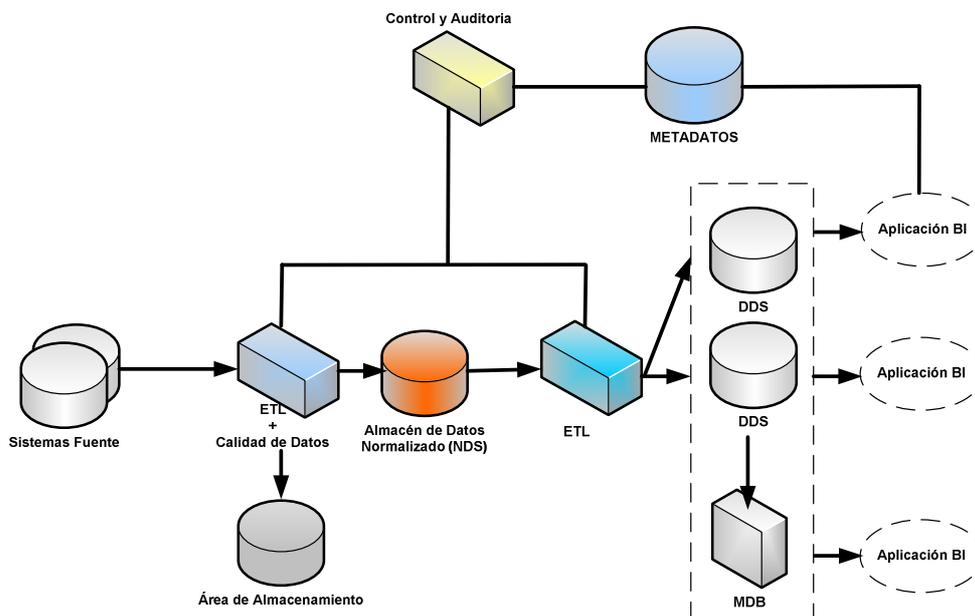


Figura 1.10 Arquitectura (NDS + DDS) [5]

1.2.4.1.4 Arquitectura con un almacén de datos operacional más un almacén de datos dimensional (ODS + DDS)

Esta arquitectura está enfocada a la creación de Data Marts operacionales (Oper Marts) ya que propone primero la construcción de un almacén de datos operacional (ODS) y luego la construcción de los almacenes de datos dimensionales que se requieran. Como ya se mencionó anteriormente un ODS

[5] RAINARDI, Vincent. **Building a Data Warehouse: With Examples in SQL Server.**

contiene únicamente los datos maestros más recientes y no almacena datos históricos.

Esta arquitectura es muy similar a la arquitectura (NDS + DDS) y se relaciona con la propuesta de Bill Inmon. Al fusionar esta arquitectura con la arquitectura (NDS + DDS) descrita anteriormente se logra tener una fábrica de información corporativa (CIF) como lo propone Inmon. Además, en ésta arquitectura el almacén de datos operacional (ODS) contiene datos limpios y de calidad lo más recientes como la tecnología lo permita por lo que éstos pueden ser consumidos por aplicaciones operacionales como sistemas OLTP, reportes operacionales o CRMs (Customer Relationship Management) [27].

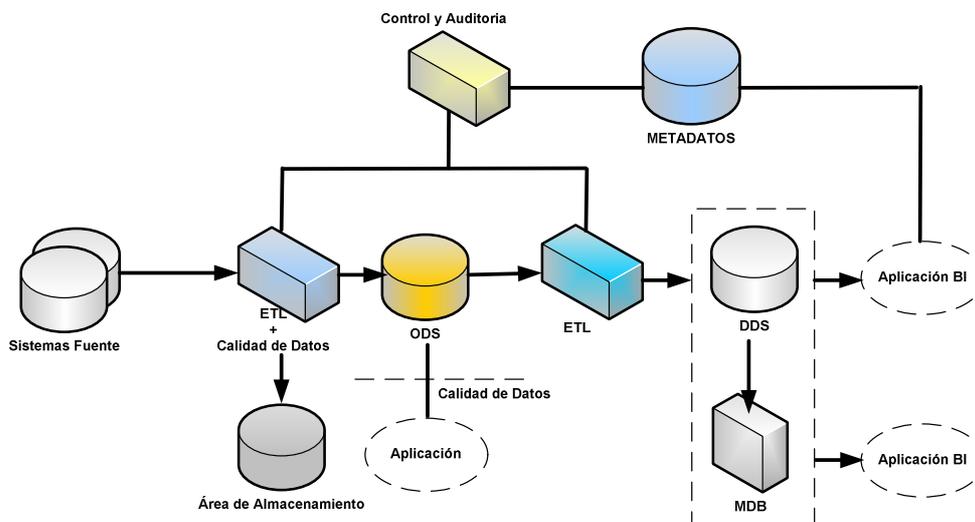


Figura 1.11 Arquitectura (ODS + DDS) [5]

1.2.4.2 Arquitectura del sistema

El diseño de la arquitectura del sistema de un Data Warehouse se refiere a la infraestructura física en dónde el sistema se ejecutará. La arquitectura física contiene tres elementos principales que son el hardware, las redes y el almacenamiento.

[5] RAINARDI, Vincent. **Building a Data Warehouse: With Examples in SQL Server.**

[27] WIKIPEDIA, The Free Encyclopedia. **CRM.**

El hardware se refiere a los servidores que usará el sistema de Data Warehouse, éstos pueden ser servidores de bases de datos, servidores de aplicaciones, servidores ETL, etc. Las redes constituyen el medio de comunicación entre los servidores y es importante que su rendimiento y disponibilidad sean altos. El almacenamiento es otro elemento importante en la arquitectura de un Data Warehouse ya que se refiere a los discos duros que se usarán, sus arreglos, configuración y tecnologías así como las soluciones de respaldos que se implantarán.

Es importante destacar que un arquitecto de Data Warehouse no diseña la arquitectura física del sistema sino sólo la arquitectura de flujo de datos, pero podría ser útil si conoce sobre los temas necesarios para su diseño. En la siguiente figura se muestra un diagrama de ejemplo de arquitectura física para un sistema de Data Warehouse.

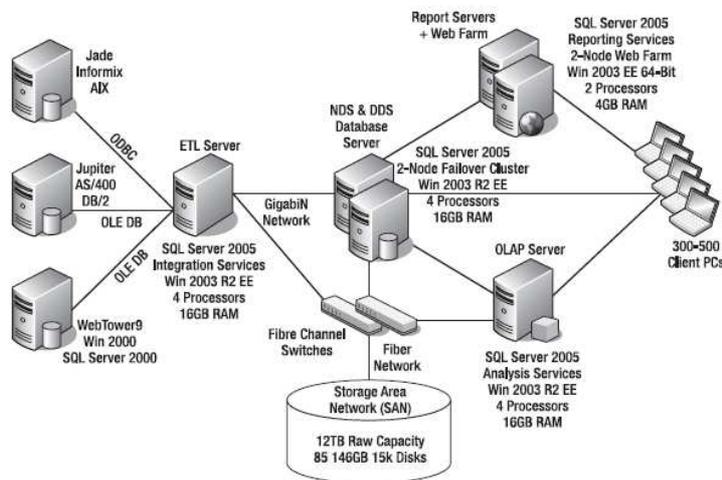


Figura 1.12 Arquitectura física de un sistema de Data Warehouse [5]

1.3 METODOLOGÍA DE DESARROLLO

Una metodología de desarrollo es un procedimiento o conjunto de procedimientos que se siguen para construir un sistema. En Ingeniería de Software, la disciplina que estudia el proceso para desarrollar sistemas de información es conocida

[5] RAINARDI, Vincent. **Building a Data Warehouse: With Examples in SQL Server.**

como el ciclo de vida del desarrollo de sistemas (SDLC, Systems Development Life Cycle) [28].

Un sistema de Data Warehouse tiene que ser construido para satisfacer las necesidades de los usuarios, el sistema tiene que ser usable.

Existen dos metodologías principales para construir un sistema de Data Warehouse, la metodología en cascada y la metodología iterativa. Inmon llama a la metodología en cascada como SDLC y a la metodología iterativa o en espiral como CLDS, lo contrario de SDLC. Además, la metodología SDLC es dirigida por requerimientos mientras que la metodología CLDS es dirigida por datos. [5]

1.3.1 METODOLOGÍA EN CASCADA

La metodología en cascada consta de ciertos pasos que deben cumplirse uno después del otro. Estos pasos en general son los siguientes:

- a) Estudio de Factibilidad,
- b) Requerimientos,
- c) Arquitectura,
- d) Diseño,
- e) Desarrollo,
- f) Pruebas,
- g) Implantación,
- h) Operación,
- i) Configuración de la Infraestructura, y
- j) Gestión del Proyecto

En la Figura 1.13 se ilustra la metodología en cascada y todos sus pasos que se deben seguir secuencialmente.

[5] RAINARDI, Vincent. **Building a Data Warehouse: With Examples in SQL Server.**

[28] WIKIPEDIA, The Free Encyclopedia. **Systems Development Life Cycle.**

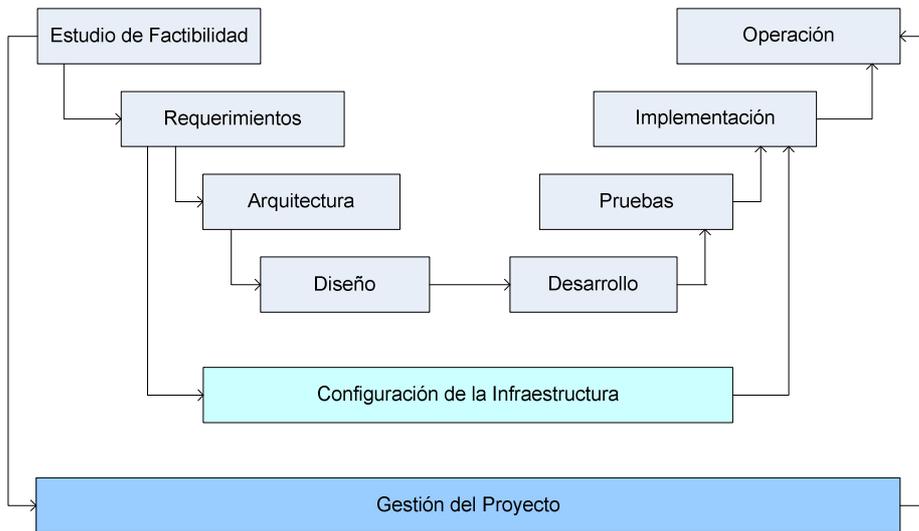


Figura 1.13 Metodología en Cascada [5]

1.3.1.1 Estudio de factibilidad

En el estudio de factibilidad se obtienen los requerimientos a alto nivel. Se responden a las preguntas ¿por qué se necesita un Data Warehouse?, ¿por qué el Data Warehouse es la solución correcta? Además, se analizan los sistemas fuente para tener una idea de la complejidad de obtención de datos. También se escribe una propuesta o caso de negocio en dónde se mencionan los beneficios, el tiempo de duración del proyecto, el costo, los requerimientos resumidos, la organización del proyecto y su plan.

1.3.1.2 Requerimientos

En la etapa de requerimientos se conversa con los usuarios para comprender los procesos del negocio, los datos, y los aspectos que serán relevantes para el desarrollo y consecución del proyecto. También se discute sobre el significado de los datos y se documentan los requerimientos funcionales y no funcionales.

1.3.1.3 Arquitectura

En esta etapa es necesario determinar la arquitectura de flujo de datos que se usará para construir el sistema de Data Warehouse. También es importante definir la arquitectura física del sistema.

1.3.1.4 Diseño

Se tiene que diseñar tres partes principales del sistema de Data Warehouse, los almacenes de datos, el sistema ETL y las aplicaciones de presentación de datos al usuario. Adicional a esto se puede diseñar los sistemas de calidad de datos y los metadatos como parte del sistema ETL.

1.3.1.5 Desarrollo

Se construyen las tres partes del sistema de Data Warehouse diseñadas en el paso anterior. Estas puede ser construidas en paralelo o definiendo interfaces funcionales entre ellas.

1.3.1.6 Pruebas

En esta etapa se prueban los almacenes de datos, los procesos ETL y las aplicaciones del usuario final. Esta etapa es crítica durante el ciclo de vida del proyecto ya que es la primera vez en que se prueban todos los componentes creados.

1.3.1.7 Implantación

Una vez que todo el sistema de Data Warehouse está listo se colocan todos los componentes en producción. El sistema ETL, los almacenes de datos y las aplicaciones de usuario comienzan a operar en un ambiente real. También se producen las guías del usuario, la guía de operación y la guía de solución de problemas comunes. Adicionalmente, se capacita a los usuarios del negocio, al equipo de operación y se cierra el proyecto.

1.3.1.8 Configuración de la infraestructura

Esta etapa consiste en crear la arquitectura física del sistema de Data Warehouse. Se realizan diseños técnicos, se compra, instala y configura el software y hardware necesario, se realizan pruebas de conectividad entre las redes y equipos con el objetivo de dejar todo listo para el equipo de operación.

1.3.1.9 Gestión del proyecto

En esta etapa se da seguimiento al plan del proyecto conociendo el estado de cada tarea, quién está haciendo qué y cuándo. También se elaboran planes de

comunicación entre los involucrados del proyecto y se busca mitigar al máximo los riesgos.

1.3.2 METODOLOGÍA ITERATIVA

El principio básico de la metodología iterativa consiste en realizar porciones pequeñas del proyecto para descubrir problemas tempranos y realizar los ajustes necesarios cuando los requerimientos no pueden ser claramente definidos por los usuarios.

Se libera una versión del sistema de Data Warehouse a producción las veces que sean necesarias para que pueda ser usada y validada por los usuarios finales. El Data Warehouse se lo construye horizontalmente siguiendo las fases de requerimientos, diseño, desarrollo, pruebas e implantación.

La metodología iterativa ayuda a minimizar los riesgos del proyecto debido a las fases incrementales que se realizan y puede ser representada como una espiral. En la Figura 1.14 se ilustra la metodología iterativa.

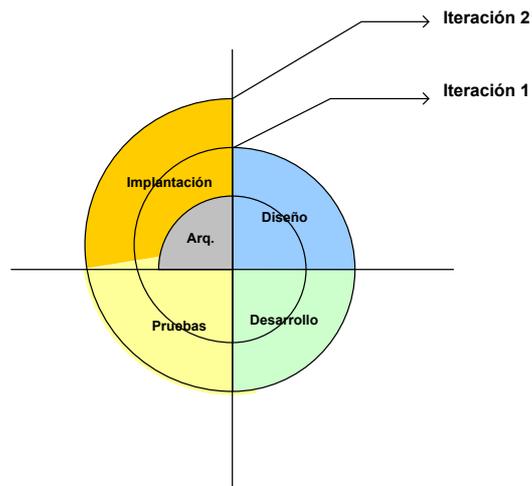


Figura 1.14 Metodología Iterativa

Según Imhoff la metodología para construir un ambiente de inteligencia de negocios es iterativa por naturaleza, y propone los siguientes pasos:

1. Seleccionar y documentar el problema del negocio ha ser resuelto a través de un Data Mart u otro.

2. Obtener tantos requerimientos como se pueda y refinarlos en los siguientes pasos.
3. Determinar la tecnología de usuario final apropiada para soportar la solución. Esta puede ser OLAP, minería de datos, aplicaciones analíticas u otras.
4. Construir un prototipo de Data Mart para probar su funcionalidad con los usuarios del negocio, rediseñándolo sí es necesario.
5. Desarrollar el modelo de datos del Data Warehouse basado en los requerimientos del usuario y en el modelo de datos del negocio.
6. Mapear los requerimientos del Data Mart al modelo de datos del Data Warehouse.
7. Generar el código para realizar los procesos ETL y los procesos de entrega de datos. Se tienen que incluir procesos de detección y corrección de errores, y procesos de auditoría.
8. Probar los procesos de creación del Data Warehouse y Data Mart. Además, medir los parámetros de calidad de datos y crear los metadatos necesarios.
9. Hasta la aceptación, mover la primera iteración del Data Warehouse y del Data Mart a producción. Capacitar a los usuarios del negocio y comenzar a planear la siguiente iteración. [4]

Es importante considerar que de ninguna forma se recomienda construir un Data Warehouse que contenga todos los datos empresariales estratégicos sin antes construir la primera capacidad analítica, el Data Mart.

[4] IMHOFF, Claudia; GALEMMO, Nicholas; GEIGER, Jonathan G. **Mastering Data Warehouse Design Relational and Dimensional Techniques.**

CAPÍTULO II

2 ANÁLISIS DEL CASO DE ESTUDIO

2.1 DESCRIPCIÓN DEL PROBLEMA

En los últimos años, la actividad aseguradora ha tenido un crecimiento considerable y ha comenzado a tomar importancia en la sociedad ecuatoriana. Hechos como la venta de vehículos nuevos con una póliza de seguro incluido y el seguro obligatorio de accidentes de tránsito (SOAT) [29] han influido en las personas para que éstas adopten una actitud de previsión ante el riesgo y reconozcan la importancia que tienen los seguros tanto para sus bienes patrimoniales como para sus vidas.

Las empresas aseguradoras pueden operar en uno o varios ramos de seguros, y cada ramo agrupa ciertos riesgos de características similares. Los ramos de seguros pueden ser patrimoniales o de daños, ramos personales, y ramos de prestación de servicios [6]. Así el seguro de vehículos cubre al vehículo ante los riesgos de robo, choque, volcadura, incendio, daños a terceros, etc. El seguro de vida cubrirá a una persona ante el riesgo de muerte, y el seguro de asistencia en viajes para vehículos cubrirá a un vehículo ante el riesgo de pérdida de equipajes o desperfectos en el funcionamiento del vehículo.

Toda actividad aseguradora se sustenta en la ciencia actuarial, la misma que estudia los principios básicos y estructurales de ésta actividad, tanto en su aspecto financiero, técnico, matemático y estadístico, para obtener un equilibrio de resultados. La ciencia actuarial se fundamenta en la ciencia estadística, la misma que utiliza, fundamentalmente, dos medios matemáticos para conseguir sus fines: la ley de los grandes números y el cálculo de probabilidades. [6]

[6] MAPFRE, Fundación. **Curso de Introducción al Seguro.**

[29] Seguro Obligatorio de Accidentes de Tránsito, **SOAT Ecuador.**

La actividad aseguradora tiene que necesariamente apoyarse en la ley de los grandes números y el cálculo de probabilidades. Por ésta razón el fin primordial de cualquier entidad aseguradora es la de conseguir un volumen de riesgos asegurados (masa asegurable) lo suficientemente amplio para dar solidez técnico-actuarial a su actividad. Esta masa asegurable tiene que hacer viable la aplicación de los principios de la ciencia actuarial y el cálculo de probabilidades.

Como consecuencia del fin primordial de la actividad aseguradora de conseguir una masa asegurable lo suficientemente amplia, las bases de datos de los sistemas de información que almacenan los datos de la masa asegurable poseen una estructura compleja sobre la cuál el realizar análisis sobre los datos resulta lento y muchas veces inútil. Debido a esto nace la necesidad de crear un almacén de datos histórico que posea un modelo de datos optimizado para el análisis, provea tiempos de respuesta óptimos, y apoye al proceso de toma de decisiones estratégicas, tácticas y operativas para el ramo de vehículos de una empresa aseguradora.

La implementación de un Data Mart que contenga datos históricos de la masa asegurable para el ramo de vehículos de una empresa aseguradora permitirá realizar las diversas actividades analíticas para conocer ciertos indicadores como el número de vehículos siniestrados, la siniestralidad, el número de vehículos asegurados, la frecuencia, causalidad y severidad de siniestros. Además, el hecho de poseer datos históricos de la masa asegurable permitirá realizar estadísticas y calcular probabilidades sobre los datos, actividades fundamentales de toda actividad aseguradora para lograr un equilibrio de resultados. La información generada por el Data Mart será presentada a través de reportes, para la toma de decisiones tácticas y operativas, y tableros de mando integral (dashboards) para la toma de decisiones estratégicas.

El Data Mart que se construirá en el proyecto de titulación será desarrollado siguiendo una metodología en cascada. Se construirá la primera capacidad analítica, el Data Mart, para la masa asegurable del ramo de vehículos de una empresa aseguradora. Se tendrá como objetivo lograr que el Data Mart desarrollado provea de las capacidades analíticas suficientes al negocio

asegurador. Todas las actividades de inteligencia de negocios (BI) que van desde la obtención de datos hasta la entrega de información a los usuarios finales estará apoyada en su totalidad por la plataforma de inteligencia de negocios BusinessObjects Enterprise XI Release 2 [10].

2.1.1 OBJETIVOS

- Aplicar el estado del arte de la inteligencia de negocios y sus herramientas para apoyar a la toma de decisiones estratégicas, tácticas y operativas del ramo de vehículos de una empresa aseguradora.
- Implementar el Data Mart para el ramo de vehículos de una empresa aseguradora.
- Presentar a través de reportes y tableros de mando integral (dashboards) la inteligencia generada por el Data Mart implementado.
- Estandarizar las actividades de inteligencia de negocios, que van desde la obtención de datos hasta la entrega de información a usuarios finales utilizando la plataforma de inteligencia de negocios BusinessObjects Enterprise XI Release 2 [10].

2.1.2 ALCANCE

En el presente proyecto de titulación se aplicará el estado del arte de la inteligencia de negocios y sus herramientas. Se analizará los requerimientos del ramo de vehículos de una empresa aseguradora para la construcción de un Data Mart que apoye a la toma de decisiones gerencial, y se obtendrá un Data Mart con datos históricos de la masa asegurable, reportes y tableros que generen inteligencia sobre el negocio.

2.2 LA EMPRESA DE SEGUROS

2.2.1 DEFINICIÓN

Según la fundación MAPFRE, una empresa de seguros o institución aseguradora puede ser definida como “la manifestación técnica y organizada de las actividades

[10] BusinessObjects, an SAP Company. **BusinessObjects XI 3.0.**

socio-económicas de compensación de riesgos. La existencia del riesgo, junto con la reparación de las consecuencias dañosas que su ocurrencia (siniestro) pueda producir, es el elemento básico que da razón de ser a la institución.”

Existen dos manifestaciones grandes de una institución aseguradora, la seguridad social y los seguros privados. Las personas sean naturales o jurídicas pueden tener diferentes actitudes frente al riesgo, éstas actitudes son la indiferencia, la prevención y la previsión.

Cuando una persona sea natural o jurídica tiene una actitud de previsión frente a un riesgo, el cuál lo transfiere a una tercera persona, conocida como la institución aseguradora, se habla de seguros. Caso contrario se habla de ahorro y auto seguro, actividades ajenas a la actividad aseguradora.

La legislación del Ecuador, en el Artículo 3 de la Ley General de Seguros [32], considera que “son empresas que realicen operaciones de seguros las compañías anónimas constituidas en el territorio nacional y las sucursales de empresas extranjeras, establecidas en el país, en concordancia con lo dispuesto en la presente ley y cuyo objeto exclusivo es el negocio de asumir directa o indirectamente o aceptar y ceder riesgos en base a primas.” Conforme a la legislación ecuatoriana las empresas aseguradoras son aquellas que realizan única y exclusivamente actividades de seguros. De igual forma, las empresas de seguros son: de seguros generales, de seguros de vida y las que operan en conjunto en las dos actividades. La Superintendencia de Bancos y Seguros es el órgano de control para las empresas de seguros privados en el Ecuador.

2.2.2 EL RIESGO

El riesgo expresa indistintamente dos ideas diferentes. El riesgo como el objeto asegurado y el riesgo como la posible ocurrencia al azar de un evento que produce una necesidad económica. La posible ocurrencia de un riesgo se previene y garantiza en la póliza de seguros, y la necesidad económica obliga al asegurador a realizar la prestación que le corresponda.

[32] Superintendencia de Bancos y Seguros – Ecuador. **Ley General de Seguros.**

El Decreto Supremo 1147 de la Legislación sobre el Contrato de Seguros [33], en su Artículo 4 define al riesgo como, “el suceso incierto que no depende exclusivamente de la voluntad del solicitante, asegurado o beneficiario, ni de la del asegurador y cuyo acontecimiento hace exigible la obligación del asegurador.”

Un riesgo para que pueda ser asegurable, tiene que cumplir con las siguientes características:

- Aleatorio o Incierto. No se conoce a ciencia cierta sí ocurrirá o no.
- Posible. El riesgo debe poder suceder.
- Concreto. Sus características y naturaleza deben ser expresadas de forma concreta.
- Lícito. No puede ir en contra de las leyes que rige la Ley General de Seguros.
- Fortuito. Es ajeno a la voluntad humana de producirlo.
- De contenido económico. Tiene que ser indemnizable económicamente.

Antes de asumir la cobertura de un riesgo, la institución aseguradora debe establecer las características del riesgo y los límites de su aceptación a través del tratamiento del riesgo, seleccionándolo, analizándolo, evaluándolo, compensándolo y distribuyéndolo.

2.2.3 EL SEGURO

El seguro es una actividad económica-financiera. Tiene como fin transformar los riesgos asegurables en pagos económicos periódicos y presupuestables.

El riesgo asegurable tiene vinculado un bien material conocido como objeto asegurado. Así, el bien material vinculado al riesgo de choque será el vehículo, y este a su vez constituirá el objeto asegurado del contrato de seguro.

[33] Superintendencia de Bancos y Seguros – Ecuador. **Legislación sobre el Contrato de Seguros - Decreto Supremo 1147.**

2.2.3.1 El Contrato de seguro

El Código Civil de la República del Ecuador [34], define en su Artículo 1454, “contrato o convención es un acto por el cual una parte se obliga para con la otra a dar, hacer o no hacer alguna cosa. Cada parte puede ser una o muchas personas.”

De acuerdo a las obligaciones a las que se comprometen ambas partes en un contrato. El contrato de seguro, “es el documento o póliza suscrita con una institución aseguradora en las que se establecen las normas que han de regular la relación contractual de aseguramiento entre asegurador y asegurado, especificándose sus derechos y obligaciones respectivas.” [6]

La buena fe tiene una especial importancia en el contrato de seguro, obligando al asegurado a describir total y claramente el riesgo, procurar evitar un siniestro y disminuir sus consecuencias en caso de producirse. De igual forma, obliga al asegurador a facilitar la información exacta de los términos del contrato y a redactar con claridad las cláusulas de las pólizas.

Los principios básicos del contrato de seguro son:

- Consensual. Es de consentimiento mutuo entre las partes.
- Bilateral. Existe una obligación recíproca entre las partes.
- Aleatorio. Existe la posibilidad de pérdida o de ganancia en una de las partes.
- Oneroso. Cada parte obtiene una prestación a cambio de otra que ha de realizar.
- De adhesión. Una parte fija las condiciones y la otra las acepta.
- Basado en la buena fe. Como lo menciona el Artículo 1562 del Código Civil [34].
- Solemne. Se cumple con las formalidades que establece la ley.

[6] MAPFRE, Fundación. **Curso de Introducción al Seguro.**

[34] LEXIS Ecuador. Legislación Indexada Sistemática. **Codificación del Código Civil de la República del Ecuador.**

Un contrato de seguro posee tres elementos fundamentales. Elementos materiales, personales y formales. Los elementos materiales corresponden al riesgo u objeto asegurado, la prima, el siniestro y la indemnización. Los elementos personales corresponden al asegurador, contratante, beneficiario, asegurado y perjudicado. Los elementos formales corresponden a la póliza, documento que formaliza el contrato de seguro.

Los elementos esenciales que todo contrato de seguros debe tener, de acuerdo al Artículo 2 del Decreto Supremo 1147 [33], son:

1. El asegurador;
2. El solicitante;
3. El interés asegurable;
4. El riesgo asegurable;
5. El monto asegurado o el límite de responsabilidad del asegurador, según el caso;
6. La prima o precio del seguro; y,
7. La obligación del asegurador, de efectuar el pago del seguro en todo o en parte; según la extensión del siniestro.

Según la Legislación del Ecuador, la falta de cualquiera de los elementos esenciales mencionados anteriormente, anulará un contrato de seguro.

2.2.3.1.1 Elementos materiales

Los elementos materiales del contrato de seguro son el objeto asegurado, la prima, el siniestro y la indemnización.

El objeto asegurado, es el objeto sobre el cual recae la cobertura garantizada por el seguro. Todo objeto asegurado tiene un capital o valor asegurado, que corresponde a la cantidad máxima que el asegurador está obligado a pagar en caso de producirse un siniestro.

[33] Superintendencia de Bancos y Seguros – Ecuador. **Legislación sobre el Contrato de Seguros - Decreto Supremo 1147.**

La prima es la aportación económica que ha de satisfacer el contratante o asegurado a la institución aseguradora por concepto de contraprestación por la cobertura del riesgo que ésta le ofrece. Según su situación contable, la prima puede ser emitida, pagada, pendiente y anulada.

El siniestro es un acontecimiento que, por causar daños previstos en la póliza, obliga a la entidad aseguradora a satisfacer al asegurado o a sus beneficiarios el capital garantizado en el contrato. Por su situación contable, los siniestros pueden ser declarados, pagados y pendientes; según el grado de intensidad del daño producido el siniestro puede ser total o parcial.

La indemnización es el importe económico que está obligado a pagar contractualmente el asegurador en caso de producirse un siniestro. Si el siniestro ha tenido carácter total, la indemnización equivale al 100% del valor asegurado. Si el siniestro ha sido parcial, se establece una proporción del valor asegurado.

2.2.3.1.2 Elementos personales

Los elementos personales del contrato de seguro son el asegurador, el contratante o solicitante, el beneficiario, el asegurado y el perjudicado.

El asegurador es la persona jurídica que mediante la formalización de un contrato de seguro, asume las consecuencias dañosas producidas por la realización de un evento cuyo riesgo es objeto de cobertura. El asegurador corresponde a la institución aseguradora.

El contratante o solicitante es la persona que suscribe una póliza y se obliga al pago de la prima. El asegurado es la persona que en sí misma o en sus bienes está expuesto el riesgo. El beneficiario es la persona designada por el asegurado o contratante como titular de los derechos indemnizatorios de una póliza. El perjudicado es la persona que, a consecuencia de un siniestro, sufre un daño o perjuicio.

En los riesgos patrimoniales, contratante, asegurado y beneficiario suelen ser la misma persona. Esto sucede generalmente en los seguros de vehículos dónde el

contratante es el propietario del vehículo, de igual forma es el asegurado y el beneficiario para recibir la indemnización necesaria si existiera un siniestro.

2.2.3.1.3 Elementos formales

La póliza es el documento que instrumenta el contrato de seguro. Este documento debe redactarse en castellano, se extenderá por duplicado y será firmado por los contratantes [33]. En la póliza se reflejan las normas que, de forma general, particular o especial regulan las relaciones contractuales convenidas entre el asegurador y asegurado.

Sólo cuando la póliza ha sido emitida y aceptada por el asegurador y asegurado tienen vigencia los derechos y obligaciones que de ella se deriven. Generalmente la vigencia de una póliza es de un año o más. Además, toda póliza de acuerdo al Decreto Supremo 1147 [33] deberá contener los siguientes datos:

- a. El nombre y domicilio del asegurador;
- b. Los nombres y domicilios del solicitante, asegurado y beneficiario;
- c. La calidad en que actúa el solicitante del seguro;
- d. La identificación precisa de la persona o cosa con respecto a la cual se contrata el seguro;
- e. La vigencia del contrato, con indicación de las fechas y horas de iniciación y vencimiento, o el modelo de determinar unas y otras;
- f. El monto asegurado o el modo de precisarlo;
- g. La prima o el modo de calcularla;
- h. La naturaleza de los riesgos tomados a su cargo por el asegurador;
- i. La fecha en que se celebra el contrato y la firma de los contratantes;
- j. Las demás cláusulas que deben figurar en la póliza de acuerdo con las disposiciones legales.

Las partes de una póliza son las condiciones generales, particulares, especiales, anexos modificatorios y documentos adicionales.

[33] Superintendencia de Bancos y Seguros – Ecuador. **Legislación sobre el Contrato de Seguros - Decreto Supremo 1147.**

- **Condiciones Generales:** conjunto de principios básicos que establece el asegurador para regular todos los contratos de seguros que se emitan en el mismo ramo. Contiene normas relativas a la extensión y objeto del seguro, riesgos incluidos con carácter general, forma de liquidación de siniestros, pago de indemnizaciones, cobro de recibos, comunicaciones mutuas entre asegurador y asegurado entre las más relevantes. Las condiciones generales para un ramo específico deben ser aprobadas por la Superintendencia de Bancos y Seguros de acuerdo al Artículo 25 de la Ley General de Seguros, y el Artículo 89 de la Legislación sobre el contrato de seguros [30].

- **Condiciones Particulares:** aspectos concretos relativos al riesgo individualizado que se asegura. Nombre y domicilio de las partes contratantes, designación del asegurado y beneficiario, concepto en el cuál se asegura, naturaleza del riesgo cubierto, designación de los objetos asegurados y su situación, valor asegurado, alcance de la cobertura, importe de la prima, recargo e impuestos, vencimiento de las primas, así como el lugar y forma de pago, duración del contrato, nombre del agente o agentes sí intervienen.

- **Condiciones Especiales:** sirven para perfilar el contenido de algunas normas recogidas en las condiciones generales y particulares. Pueden encontrarse impresas como anexos, para ser incorporadas a una póliza.

- **Anexos Modificatorios:** tienen como objetivo modificar aspectos substanciales de forma o de fondo de las pólizas.

- **Documentos Adicionales:** corresponden a todos los documentos que se incorporan en la póliza. Estos son la solicitud, descripción del objeto a asegurar para su determinación técnica del riesgo, y la inspección del riesgo.

Según Meythaler, todo lo que constituye una póliza es la solicitud del tomador, la aceptación del asegurador, la inspección del riesgo, la póliza misma con sus condiciones generales, particulares y especiales, todos los anexos de distinta índole y hasta la correspondencia cruzada entre los contratantes.

[30] MEYTHALER, Eduardo. *El Contrato Partes y Elementos – Aspectos Legales*.

Una póliza según la amplitud de riesgos que cubra podrá ser simple, combinada y a todo riesgo. La póliza simple cubre una garantía concreta. La póliza combinada cubre diferentes riesgos, y la póliza a todo riesgo garantiza conjunta y simultáneamente todos los riesgos que puedan afectarle a un objeto asegurado. En el seguro de vehículos la póliza a todo riesgo implica la cobertura de todos los riesgos habituales: daños del vehículo, responsabilidad civil, robo, choque, etc.

2.2.4 EL REASEGURO

El reaseguro corresponde a una técnica de distribución de riesgos mediante la cual un asegurador cede todo o parte de un riesgo a una o varias empresas conocidas como reaseguradores, de tal forma de que si ocurriese un siniestro que afecte a su cartera, el asegurador no se vea afectado económicamente al indemnizar la totalidad de un siniestro, sino sólo parte de éste.

El Artículo 85 del Decreto Supremo 1147 define, “el reaseguro es una operación mediante la cual el asegurador cede al reasegurador la totalidad o una parte de los riesgos asumidos directamente por él. Reasegurador es la persona o entidad que otorga el reaseguro; puede también llamarse cesionario o aceptante.”

Un contrato de reaseguros es independiente del contrato de seguro, el contrato de reaseguros no modifica las obligaciones asumidas por el asegurador. Además, no da al asegurado acción directa contra el reasegurador. Es decir, el reasegurador es transparente para el asegurado, y es el asegurador el único encargado de hacer cumplir las obligaciones del reasegurador.

2.2.5 CLASES DE SEGUROS

Los seguros pueden ser clasificados en base a la naturaleza de los riesgos en seguros de personas, seguros de daños o patrimoniales, y en seguros de prestación de servicios.

- Seguros de personas. El objeto asegurado es la persona dependiendo de su existencia, salud o integridad el pago de la prestación. Las principales modalidades de ésta clasificación de seguros son el seguro de vida, el seguro de accidentes y el seguro de enfermedad.

- Seguro de daños o patrimoniales. Bajo esta clasificación se agrupan todos los seguros cuyo fin es el de reparar el patrimonio del contratante del seguro por una pérdida causada en un siniestro. Se pueden dividir en seguros de cosas o bienes materiales y seguros de responsabilidad civil.
- Seguro de prestación de servicios. El asegurador garantiza la prestación de un servicio cuando se produzca una situación que lo haga necesario, como la defensa jurídica, asistencia médica, sepelio, repatriación de vehículos, etc.

2.2.6 ORGANIZACIÓN DE UNA EMPRESA DE SEGUROS

Una empresa de seguros como toda empresa se constituye por un conjunto de personas que bajo la autoridad de un director general poseen las competencias necesarias para conocer, debatir y resolver ciertos asuntos y materias propias de la actividad aseguradora. En una empresa de seguros existe la gestión comercial, técnica, financiera y de servicios generales.

En la Figura 2.1 se ilustra un organigrama ideal para una empresa de seguros; ideal debido a que éste depende del equipo humano y el volumen de la empresa. El organigrama de la Figura 2.1 considera los siguientes órganos. El órgano de gestión comercial que se encarga de las actividades comerciales de seguros. El órgano de gestión técnica que se encarga de analizar los resultados técnicos de la empresa con el objetivo de lograr un equilibrio de resultados. El órgano de gestión financiera que se encarga del manejo de inversiones y el control contable de las operaciones económico-financieras, y el órgano de servicios generales que tiene como objetivo proveer diversos servicios para mejorar la atención a los clientes.

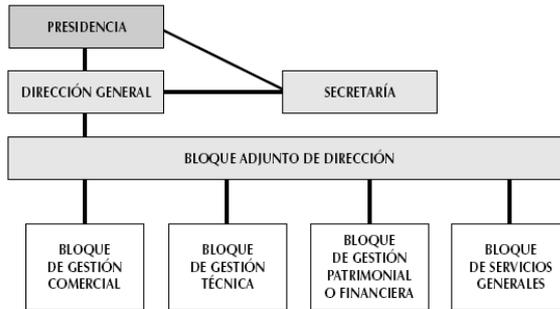


Figura 2.1 Organigrama ideal de una empresa de seguros [6]

El modelo de organización de una empresa de seguros depende de los siguientes factores:

- Organización Directiva. Organización de los órganos socioempresariales, toman decisiones referentes a la dirección, gestión y administración de la empresa. Las principales manifestaciones de estos órganos son la asamblea general, el consejo de administración, la comisión ejecutiva, el consejero delegado y el director general o gerente.
- Política de la empresa. La organización de la empresa de seguros depende de los objetivos que se quieren conseguir a corto, mediano y largo plazo, y de la manera de lograr su consecución.
- Volumen de operaciones. Es un factor principal en la organización de una empresa de seguros debido a que a mayor tamaño de la empresa, mayor es la necesidad de dividir funciones.
- Clases de operaciones. De acuerdo a las operaciones de seguros, la organización puede variar. Por ejemplo, el ramo de incendios permite gran centralización en la prestación de servicios. Al contrario, el ramo de vehículos exige servicios autónomos y extensos debido a la movilidad del riesgo.
- Medios de procesamiento de datos. Influyen en la estructura de los departamentos y servicios, así como también en la necesidad de dividir ciertas funciones.

[6] MAPFRE, Fundación. **Curso de Introducción al Seguro.**

2.2.6.1 Organización funcional de una empresa de seguros

Corresponde a la organización operativa de la empresa de seguros. Depende de la actividad que desarrolla la empresa, el entorno en el que opere y de la política de gestión que se practique. La organización funcional de la empresa de seguros puede ser según el grado de centralización, la ubicación geográfica de la estructura administrativa y el grado de agrupación de los servicios de idéntica naturaleza.

2.2.6.1.1 Según el grado de centralización

Según el grado de centralización. La organización puede ser centralizada o descentralizada.

- Organización centralizada. Todas las responsabilidades de actuación y resultados recaen sobre los órganos centrales. Sí existen órganos territoriales o periféricos, éstos no pueden tomar decisiones o pueden hacerlo de forma muy limitada.
- Organización descentralizada. Busca conseguir mayor flexibilidad, rapidez y eficacia en la prestación de servicios. Los órganos territoriales tienen responsabilidades de actuación y resultados.

2.2.6.1.2 Según la ubicación geográfica de la estructura administrativa

Según la ubicación geográfica de la estructura administrativa. La organización puede ser central o territorial.

- Organización central. La oficina principal ejerce todas las funciones directivas. La organización territorial se constituye por agencias, las cuales únicamente ejercen funciones administrativas y de tramitación.
- Organización territorial. Se constituye por sucursales dedicadas exclusivamente a la gestión comercial y a las operaciones derivadas de los contratos de seguros. Los órganos centrales delegan responsabilidades a los órganos territoriales para que ejerzan funciones de control, revisión y administración.

2.2.6.1.3 Según el grado de agrupación de servicios de idéntica naturaleza

Según el grado de agrupación de servicios de idéntica naturaleza. La organización puede ser vertical u horizontal.

- Organización vertical. La gestión se realiza independientemente en cada ramo en que opera la empresa. Por cada ramo existen tantos departamentos como sean necesarios. En la Figura 2.2 se ilustra este tipo de organización.

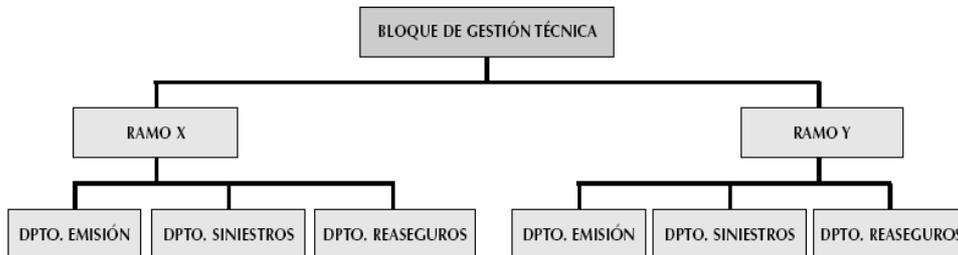


Figura 2.2 Organización vertical de una empresa de seguros [6]

- Organización horizontal. Todos los servicios comunes para todos los ramos se agrupan en un solo departamento. Es decir, cada departamento agrupa uno o más ramos. En la Figura 2.3 se ilustra este tipo de organización.

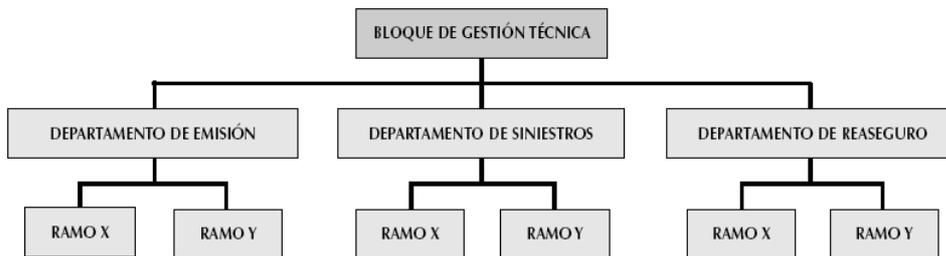


Figura 2.3 Organización horizontal de una empresa de seguros [6]

2.2.6.2 Órganos de dirección y departamentos de una empresa de seguros

Los órganos de dirección están encargados del ejercicio y control de actividades ejecutivas. Estos órganos de dirección son:

[6] MAPFRE, Fundación. **Curso de Introducción al Seguro.**

- La dirección general. Toma decisiones válidas en todas las materias.
- La dirección administrativa. Se encarga del ejercicio, desarrollo y supervisión de las funciones administrativas.
- La dirección comercial. Se encarga del ejercicio, desarrollo y supervisión de las actividades relaciones con la política y función comercial.
- La dirección financiera. Se encarga del ejercicio, desarrollo, y supervisión de las actividades relacionadas con las políticas de inversiones y contabilidad.
- La dirección técnica. Se encarga del ejercicio, desarrollo y supervisión de las actividades relacionadas con la contratación de pólizas y la tramitación de siniestros.
- La dirección de recursos humanos. Se encarga de ejercitar, desarrollar y supervisar las funciones relacionadas con las relaciones laborales, formación y comunicación del personal.

Los departamentos son órganos encuadrados en una dirección. Cada departamento ejecuta y controla actividades determinadas por cada dirección. Los principales departamentos de una empresa de seguros son:

- Departamento de contratación. Se encarga en general de la suscripción y formalización de pólizas.
- Departamento de emisión de pólizas. Se encarga de la confección y envío de pólizas.
- Departamento de cobranza. Se encarga del cobro de recibos de nueva producción o de cartera.
- Departamento de siniestros. Se encarga de la aceptación, tramitación y liquidación de siniestros.
- Departamento comercial. Se encarga del ejercicio, desarrollo y control de la actividad comercial de seguros a través de su personal vendedor.
- Departamento de inversiones. Se encarga de colocar los recursos económicos de la empresa en mercados mobiliarios o inmobiliarios.
- Departamento de contabilidad. Se encarga del control contable de las operaciones económico-financieras. Entre sus principales funciones están el llevar la contabilidad general de la empresa y pagar los impuestos.
- Departamento de informática. Se encarga del procesamiento de datos y de generar soluciones tecnológicas que apoyen a la empresa.

- Departamento de personal. Se encarga de regular las relaciones laborales con los empleados.
- Departamento de formación. Se encarga de la capacitación técnica y profesional de empleados y colaboradores de la empresa.

2.2.6.3 Proyección de una empresa de seguros hacia el futuro

En la actividad aseguradora, la necesidad de prolongar su actividad en el tiempo es un requisito necesario. Para que una empresa de seguros pueda perdurar en el tiempo es necesario que se apoye en una política comercial, en el establecimiento de objetivos y en la función comercial y estrategia empresarial.

2.2.6.3.1 Política comercial

La política comercial es un conjunto de principios que orientan a la empresa aseguradora a una determinada forma de ser frente al público. La política comercial busca aumentar el número de riesgos asegurados siempre y cuando se cumplan las siguientes condiciones:

- La política comercial tiene que estar coordinada con los objetivos de la política general de la empresa.
- La política comercial tiene que ser lo suficientemente flexible para adaptarse a las conveniencias y posibilidades de cada momento.
- La política comercial debe buscar una proyección temporal.
- La política comercial debe estar acorde a los fines que se pretenden conseguir y a los recursos con los que se cuenta.

2.2.6.3.2 Establecimiento de objetivos

Un objetivo es el fin básico que se persigue en cualquier actividad comercial. Los objetivos pueden ser objetivos de resultados, objetivos de actuación y otros objetivos que cada empresa se plantee.

Los objetivos de resultados tienen una finalidad continua. Por ejemplo, mantener una siniestralidad de toda la cartera de vehículos asegurados que esté bajo el 55%. También podría ser, el mantener un volumen de vehículos asegurados equivalente al 45% de todo el parque nacional de vehículos.

Los objetivos de actuación no son continuos y se busca alcanzarlos en un determinado plazo. Los objetivos de actuación son los medios para conseguir un objetivo de resultados. Por ejemplo, para el objetivo de resultados descrito anteriormente sobre la siniestralidad. El objetivo de actuación constituiría en realizar un estudio de siniestros, sus causas, frecuencia e intensidad de los tipos de riesgos más peligrosos que ocasionan que los vehículos se accidenten para decidir a futuro su aceptación o rechazo.

2.2.6.3.3 Función comercial y estrategia empresarial

La función comercial busca incrementar en número y cuantía sus riesgos asegurados. Los objetivos específicos de la función comercial deben ir encaminados a conseguir tanto el aumento como la conservación de la cartera. Entre los objetivos específicos de la función comercial en el seguro, están:

- Adquirir nuevas pólizas. No sólo de nuevos clientes, sino de clientes ya existentes ofreciéndoles cobertura en otros riesgos.
- Aumentar las coberturas en las pólizas existentes. Aumentando el capital asegurado o el número de coberturas de una misma póliza.
- Conservar los asegurados y sus coberturas. Evitar que se anulen las pólizas y coberturas existentes.
- Mantener buenas relaciones con los asegurados. Ayuda al incremento de la cartera y la conservación de la misma.

La estrategia empresarial de una empresa de seguros se constituye a través de la creación de un plan básico operativo que le permita llegar a la consecución de los objetivos propuestos. La estrategia empresarial tiene que ser lo suficientemente flexible para adaptar la actividad aseguradora a nuevos retos y circunstancias del mercado.

2.3 EL RAMO DE VEHÍCULOS Y LOS MEDIOS TÉCNICOS PARA SU GESTIÓN

2.3.1 LOS RAMOS Y SU CLASIFICACIÓN

Se designa como ramo al conjunto de riesgos de características similares o naturaleza semejante. Así se habla de ramo de vehículos, ramo de vida, ramo de incendios, ramo de fianzas, etc. La clasificación en ramos permite establecer lo siguiente:

- Grupos homogéneos de riesgos con características comunes,
- Un adecuado tratamiento de los riesgos y su análisis estadístico, y
- Una correcta tarificación de los riesgos.

Los ramos pueden ser clasificados de la misma forma que la clasificación de los seguros. Así se tienen ramos patrimoniales o de daños, ramos personales y ramos de prestación de servicios. [6]

Entre los ramos patrimoniales o de daños se tiene el ramo de incendios, responsabilidad civil, vehículos, agrícola, ganadero, fianzas, robo, crédito y caución, transporte, y ramo de ingeniería. El ramo de vehículos objeto de estudio en este proyecto de titulación indemniza a los beneficiarios del seguro por los accidentes producidos a consecuencia del uso y circulación de vehículos.

Entre los ramos personales se tiene el ramo de vida, el ramo de accidentes personales y el ramo de enfermedad.

Entre los ramos de prestación de servicios se tiene el ramo de asistencia médica o sanitaria, el ramo de enterramiento o deceso, el ramo de defensa jurídica, el ramo de asistencia en viaje y el ramo de seguro turístico.

[6] MAPFRE, Fundación. **Curso de Introducción al Seguro.**

2.3.2 EL SEGURO DE VEHÍCULOS

El seguro de vehículos tiene por objeto la prestación de indemnizaciones derivadas de accidentes producidos a consecuencia de la circulación de vehículos. El seguro de vehículos puede ser de suscripción obligatoria o de suscripción voluntaria.

El seguro obligatorio de vehículos está destinado a cubrir las lesiones corporales o daños materiales que el conductor de un vehículo en circulación ocasione a terceros. El seguro obligatorio en el Ecuador, conocido como SOAT (Seguro Obligatorio de Accidentes de Tránsito) [29] y cuyo reglamento fue aprobado con Decreto Ejecutivo 809 el 19 de diciembre del 2007 y publicado en el Registro Oficial 243 del 2 de enero del 2008, “ampara a cualquier persona que sufra lesiones corporales, funcionales u orgánicas, o falleciere a causa o como consecuencia de un accidente de tránsito en que se halle involucrado uno o más vehículos a motor.”

El seguro voluntario de vehículos es de libre contratación, cubre el exceso de los límites indemnizatorios del seguro obligatorio u otras garantías. Entre las coberturas principales del seguro voluntario de vehículos se tiene la responsabilidad civil para garantizar la responsabilidad en que pueda incurrir un conductor por daños materiales ocasionados a terceros, o para garantizar el exceso en la indemnización de las coberturas del seguro obligatorio para los daños corporales ocasionados a terceros. El seguro voluntario también provee de coberturas contra daño, incendio o robo del vehículo.

Es importante considerar que las indemnizaciones del seguro de vehículos se realizarán siempre y cuando el vehículo circule en vías, carreteras, parqueaderos, y demás caminos aptos para su circulación. Se excluirán entonces las competencias deportivas de vehículos, la utilización del vehículo para tareas industriales o agrícolas, y la utilización del vehículo para cometer delitos. [6][31]

[6] MAPFRE, Fundación. **Curso de Introducción al Seguro.**

[29] Seguro Obligatorio de Accidentes de Tránsito, **SOAT Ecuador.**

[31] TARRÉ, Julio. Seguros Generales – Vehículos.

2.3.2.1 Definiciones Generales

2.3.2.1.1 Automotor

Un automotor, según Tarré, es “todo aparato provisto de un motor propulsor destinado a circular por el suelo para transporte de personas o de bienes, incluyendo cualquier elemento montado sobre ruedas que pueda ser acoplado.”

Dentro de ésta definición, no quedan incluidos los vehículos que circulan sobre rieles ni los vehículos agrícolas ni industriales.

Los automotores o vehículos propulsados por un motor constituyen los objetos que más se aseguran en el ramo de vehículos, considerando además que únicamente incluyen dentro de ésta categoría los siguientes tipos de vehículos.

2.3.2.1.2 Tipos de vehículos

- Vehículos privados livianos. Estos vehículos soportan una carga útil de hasta tres toneladas y pueden llevar consigo hasta 10 pasajeros. Dentro de este tipo de vehículos se encuentran los vehículos de uso doméstico, mediana y frecuente movilización y flotas que contengan más de seis automóviles.
- Vehículos privados para transporte. Estos vehículos soportan más de tres toneladas de carga, y pueden llevar consigo más de 10 pasajeros. Generalmente son usados para transporte de carga y pasajeros tanto en zonas urbanas como rurales.
- Vehículos de transporte público. Incluyen vehículos livianos, medianos y pesados de uso público destinados para el transporte de carga y pasajeros.
- Vehículos Estatales. Son de propiedad del Gobierno y son usados para transporte de carga o movilización de pasajeros.
- Vehículos Diplomáticos. Son usados por entidades diplomáticas.
- Motos. Generalmente son de uso privado y sirven para realizar actividades de mensajería.

2.3.2.2 Póliza de vehículos

La póliza de vehículos es el instrumento legal y formal que sella las relaciones entre el asegurador y el asegurado para uno o varios vehículos que sean objeto

[31] TARRÉ, Julio. Seguros Generales – Vehículos.

de las coberturas contratadas. Para la contratación de una póliza de vehículos influyen los siguientes factores [31]:

- Tipo de vehículo.
- Año de fabricación del vehículo.
- Tipo de uso de vehículo. Sí es de uso personal, familiar, empresarial, transporte o carga.
- La entidad que paga el seguro. Sí es un individuo o una empresa.
- Las experiencias siniestralas del vehículo. Es decir los antecedentes del vehículo, para saber sí el vehículo tuvo en el pasado siniestros o no.
- La profesión, capacidad económica y estabilidad del asegurado.
- El número de personas que usan el vehículo.
- La ubicación en la cuál el vehículo permanece en las noches.
- La ubicación en la cuál permanece el vehículo durante las horas de trabajo.
- La vigencia que se quiera contratar la póliza.
- Las coberturas solicitadas.
- La facilidad y costo de reparación del vehículo.
- El lugar de operación o rodaje del vehículo.
- El valor asegurado o suma asegurada del vehículo.
- El estado del vehículo.
- La situación atmosférica del lugar en donde el vehículo se movilizará con más frecuencia.

2.3.2.2.1 Componentes de la póliza de vehículos

Una póliza de vehículos contiene los siguientes componentes [31]:

1. Número de la Póliza.
2. Nombre y dirección de la aseguradora.
3. El título de la póliza. El cuál será: "Póliza de Seguro de Vehículos"
4. Declaraciones.
5. Las coberturas o convenios de seguros.
6. Los riesgos excluidos relacionados con cada cobertura.
7. Los riesgos excluidos a toda forma de cobertura.

[31] TARRÉ, Julio. Seguros Generales – Vehículos.

8. Las condiciones generales y especiales.
9. La forma en que se realizará el arreglo de pérdidas.
10. Arbitraje.
11. Defensa de juicio.
12. Subrogación.
13. Terminación anticipada del seguro.
14. Prescripción liberatoria.
15. Jurisdicción.
16. Número de resolución de aprobación del formulario por la Intendencia General de Seguros de la Superintendencia de Bancos.
17. El cuadro de declaraciones.
18. Las condiciones particulares.
19. Las firmas de los contratantes, asegurador y asegurado.

2.3.2.2.2 Principales coberturas de la póliza de vehículos

- Responsabilidad Civil. Cobertura por daños emergentes causados involuntariamente a la persona o propiedad de terceros como consecuencia de un evento dependiente de la circulación y uso del vehículo.
- Choques y Volcaduras. Daños que sufra el vehículo asegurado como consecuencia de choques o volcaduras. El choque o volcadura puede ser ocasionado por el encuentro repentino y violento del vehículo con otro objeto, persona o animal.
- Robo Total. Cubre el riesgo de robo total del vehículo y sus consecuentes daños o pérdidas, sujeto a las siguientes condiciones. Primero, que hayan transcurrido por lo menos 60 días del reclamo y no se haya recuperado el vehículo. Por último, que la obligación del asegurador se cumpla mediante la indemnización en efectivo o con la entrega de otro vehículo de igual marca, tipo o valor equivalente.
- Incendio. Cubre el riesgo de incendio. La garantía del asegurador comprende únicamente la reparación o indemnización de los daños o pérdidas ocasionadas al vehículo asegurado.
- Rotura de vidrios. Ampara la rotura de vidrios del vehículo.
- Accidentes personales para ocupantes. Garantiza la indemnización al conductor y a los ocupantes del vehículo que cause muerte, invalidez total y

permanente, y gastos médicos para un número máximo de ocupantes asegurados.

2.3.2.2.3 Principales exclusiones de la póliza de vehículos

La póliza no ampara pérdidas o daños causados al vehículo asegurado cuando exista una o varias de las siguientes circunstancias:

- Cualquier accidente ocurrido fuera del radio de circulación amparado.
- Tránsito por carreteras o caminos no entregados oficialmente al tráfico público.
- Accidente ocurrido por violación o inobservancia de las leyes de tránsito.
- Cuando el vehículo se dañe por circulación por vías no asfaltadas o pavimentadas.
- Tránsito por riachuelos, esteros y playas.
- Cuando el vehículo sea utilizado para carreras, competencias, pruebas de velocidad, oficiales o no.
- Daños o pérdidas de los objetos transportados que no formen parte del mismo.
- Accidentes debidos a fallas o daños mecánicos, no se cubrirá la causa del siniestro, pero sí los daños consecuenciales.
- Mientras se emplee el vehículo en cualquier negocio o transporte ilícito.
- Granizo, inundación, terremoto, temblor, erupción volcánica, tifón, huracán, tornado u otra convulsión de la naturaleza.
- Los daños que sufre el vehículo por haberse puesto en marcha después de ocurrido el accidente.

De igual forma, el asegurado perderá el derecho a cualquier indemnización en cualquiera de los siguientes casos:

- Sí el interés del asegurado sobre el vehículo fuere diferente al declarado.
- Sí el vehículo es usado para otros fines no establecidos en la póliza.
- Cuando el asegurado se encuentre bajo la influencia de bebidas alcohólicas, drogas tóxicas, o estupefacientes.
- Cuando el asegurado no posea la debida credencial de manejo.
- Cuando el asegurado por negligencia dejara agravar los daños causados por el siniestro.
- Sí el asegurado no mantiene el vehículo en buen estado de conservación.

- Si el asegurado de mala fe, hiciera falsas declaraciones del monto los daños, ocultase piezas salvadas, o trate de obtener ventajas ilícitas del seguro.

2.3.3 LOS MEDIOS TÉCNICOS DEL SEGURO

Los medios técnicos del seguro permiten a las entidades aseguradoras proyectarse hacia el futuro con las debidas garantías de permanencia, equilibrio, estabilidad y solvencia. [6]

Se destacan los siguientes medios técnicos del seguro:

- La ciencia actuarial. Estadística y matemática para definir las bases técnicas y las tarifas de pólizas.
- Las provisiones técnicas o reservas, margen de solvencia y el fondo de garantía.
- Los sistemas de distribución de riesgos, el reaseguro y coaseguro. En el reaseguro el asegurador directo cede parte de los riesgos a otras aseguradoras. El coaseguro se da cuando existe concurrencia de dos o más aseguradoras en la cobertura de un mismo riesgo.
- Los instrumentos de análisis técnico. Entre los cuáles se destacan, el resultado técnico, el índice de frecuencia, el índice de intensidad, el índice de siniestralidad, la siniestralidad, la siniestralidad esperada, y las provisiones técnicas o reservas.

Los instrumentos de análisis técnico son de principal interés para el desarrollo del presente proyecto de titulación debido a que serán implementados en el Data Mart para el apoyo a la toma de decisiones y la realización de actividades analíticas. A continuación se explica cada instrumento de análisis técnico:

2.3.3.1 Resultado técnico

El resultado técnico, en una empresa de seguros, corresponde a la diferencia entre las primas recaudadas y los gastos ocasionados por siniestros pagados o

pendientes de pago. El resultado técnico considera única y exclusivamente los ingresos y gastos provenientes de la actividad aseguradora.

$$\text{Resultado Técnico} = \text{Ingresos} - \text{Gastos}$$

- **Ingresos** = Comisiones recibidas + Recobros + Primas recaudadas + Otros
- **Gastos** = Comisiones pagadas + Gastos Administrativos + Primas cedidas + Otros Gastos

Figura 2.4 Fórmula de cálculo del resultado técnico

2.3.3.2 Índice de frecuencia

El índice de frecuencia es la cifra o coeficiente que refleja el promedio de número de siniestros que una póliza o conjunto de pólizas tiene durante un año completo.

$$\text{Índice de Frecuencia} = \text{Nro. Siniestros Pagados} / 12$$

Comentario [MM2]:

Comentario [MM3]: Es importante incluir la fórmula de cálculo del índice de frecuencia acorde con la descripción.

Figura 2.5 Fórmula de cálculo del índice de frecuencia para un año completo

El índice de frecuencia también corresponde al porcentaje resultante de dividir el número de siniestros tramitados sobre el número de vehículos asegurados.

$$\text{Índice de Frecuencia} = \text{Nro. Siniestros Pagados} / \text{Nro. Objetos Asegurados}$$

Figura 2.6 Fórmula de cálculo del índice de frecuencia en base a la cartera de vehículos asegurados

2.3.3.3 Índice de intensidad

El índice de intensidad refleja el coste promedio de los siniestros producidos respecto a una póliza o conjunto de pólizas. El índice de intensidad también se lo conoce como índice de siniestro promedio y resulta de dividir entre el valor pagado en siniestros sobre el número de siniestros pagados.

$$\text{Índice de Intensidad} = \text{Siniestros Pagados} / \text{Nro. Siniestros Pagados}$$

Figura 2.7 Fórmula de cálculo del índice de intensidad o índice de siniestro promedio

2.3.3.4 Índice de siniestralidad

El índice de siniestralidad corresponde al porcentaje que refleja la proporción existente entre el coste de los siniestros producidos por una póliza o conjunto de pólizas y el volumen global de las primas que han devengado en el mismo período tales operaciones.

$$\text{Índice de Siniestralidad} = \text{Siniestros} / \text{Primas}$$

Figura 2.8 Fórmula de cálculo del índice de siniestralidad

2.3.3.5 Siniestralidad

La siniestralidad equivale a la proporción entre el importe total de los siniestros y las primas recaudadas por una entidad aseguradora. Esta proporción se mide mediante el índice de siniestralidad. La siniestralidad esperada corresponde al importe de los siniestros que, en base a experiencias anteriores, se calcula que deberá ser satisfecho.

2.3.3.6 Provisiones técnicas o reservas

Las provisiones técnicas o reservas son aquellas reservas económicas que toda entidad aseguradora debe realizar para hacer frente a obligaciones futuras. Toda entidad aseguradora tiene la obligación de constituir y mantener en todo momento provisiones técnicas suficientes para el conjunto de sus actividades. Estas provisiones tienen que estar calculadas adecuadamente, contabilizadas e invertidas en activos aptos para su cobertura.

2.3.3.7 Otros indicadores de análisis

Existen otros indicadores importantes para la gestión de seguros, estos son el valor asegurado promedio, la tasa promedio de la cartera de póliza, el porcentaje de comisiones pagadas y el porcentaje de gastos administrativos. A continuación en la Figura 2.9, Figura 2.10, Figura 2.11, y Figura 2.12 se indican las fórmulas de cálculo para dichos indicadores:

$$\text{Valor Asegurado Promedio} = \text{Suma Asegurada de la cartera} / \text{Nro. Objetos asegurados}$$

Figura 2.9 Fórmula de cálculo del valor asegurado promedio

El valor asegurado promedio indica el valor promedio por el cuál se está asegurando cada objeto asegurado de todos o un conjunto de pólizas.

$$\text{Tasa Promedio de la Cartera} = \text{Prima Neta Anual} / \text{Suma asegurada de la cartera}$$

Figura 2.10 Fórmula de cálculo de la tasa promedio de la cartera

La tasa promedio de la cartera indica el porcentaje entre la prima neta anual y la suma asegurada de la cartera. Como se puede apreciar a mayor tasa mayor será la prima neta que se tendrá que cobrar por póliza.

$$\% \text{ Comisiones Pagadas} = \text{Comisiones Pagadas} / \text{Prima Neta Anual}$$

Figura 2.11 Fórmula de cálculo del % de comisiones pagadas

El porcentaje de comisiones pagadas indica que porcentaje de las primas se han pagado por concepto de comisiones a los agentes o intermediarios de seguros. Un agente es un profesional del seguro que realiza la actividad de mediación de seguros entre una o varias entidades aseguradoras. Los agentes constituyen la prolongación de las entidades aseguradoras [6].

$$\% \text{ Gastos Administrativos} = \text{Gastos Administrativos} / \text{Prima Neta Anual}$$

Figura 2.12 Fórmula de cálculo del % de gastos administrativos

El porcentaje de gastos administrativos indica el porcentaje de las primas que se han consumido para cubrir con los gastos de administración.

Comentario [MM4]: Considerar incluir algunos de los índices mencionados en el documento: Informe Primario Sobre Índices de Seguros de No Vida Para Supervisores de Seguros.pdf. Realizar una lectura completa del capítulo para revisar ortografía, composición y estructura.

[6] MAPFRE, Fundación. Curso de Introducción al Seguro.

CAPÍTULO III

3 DESARROLLO DEL DATA MART

El capítulo de desarrollo del Data Mart para el ramo de vehículos constituye una aplicación del estado del arte de la inteligencia de negocios. En este capítulo se aplican los conceptos introducidos en el capítulo I para dar solución al problema descrito en el caso de estudio del capítulo II.

De acuerdo a lo mencionado en el capítulo I, existen dos arquitecturas de inteligencia de negocios propuestas por los dos representantes más importantes del área. La primera, conocida como CIF (Corporate Information Factory) y propuesta por Bill Inmon, propone que primero se construya el Data Warehouse y luego se deriven los Data Marts que sean necesarios. La segunda, conocida como MD (Multi Dimensional) y propuesta por Ralph Kimball, propone que primero se construya uno o los Data Marts necesarios hasta lograr un conjunto de Data Marts que constituirán el Data Warehouse para una compañía.

Para el desarrollo del Data Mart para el ramo de vehículos se empleará la arquitectura propuesta por Ralph Kimball, comenzando a construir la primera capacidad analítica conocida como Data Mart. Además, se utilizará la técnica de modelamiento dimensional para diseñar un modelo optimizado para el análisis que tenga la capacidad de responder a las diversas preguntas del negocio.

El desarrollo del Data Mart se lo realizará siguiendo una metodología en cascada. Aunque, es importante destacar que el desarrollo de un Data Mart es iterativo por naturaleza. En un ambiente de producción un Data Mart debe ser lo suficientemente flexible para responder a los cambios de requerimientos del negocio. Sí se lo desarrollaría de una forma secuencial y en cascada se anularía la característica de flexibilidad, ya que el Data Mart estaría limitado a los requerimientos iniciales definidos al inicio del desarrollo.

Se justifica el uso de la metodología en cascada para el desarrollo de éste proyecto de titulación debido a que los requerimientos serán definidos una única vez y serán usados con el propósito de explicar como seguir la metodología de desarrollo para construir el Data Mart y dar solución al problema planteado.

3.1 PLANEACIÓN

Es importante planificar las actividades que se realizarán en cada etapa del proceso de construcción del Data Mart. En éste subcapítulo se definen qué actividades se realizarán en cada fase, las herramientas de la plataforma de inteligencia de negocios que se utilizarán y los recursos con los que se cuenta para finalizar con éxito la construcción del Data Mart.

3.1.1 HERRAMIENTAS

Como ya se mencionó en los capítulos anteriores, para el desarrollo del Data Mart se utilizará la plataforma de inteligencia de negocios BusinessObjects Enterprise XI Release 2 Service Pack 3 (BOE XI R2 SP3). La plataforma contiene todas las herramientas necesarias para adoptar una arquitectura de inteligencia de negocios sea ésta CIF o MD, apoyando desde los procesos de carga de datos hasta la entrega de información a los usuarios finales a través de reportes o tableros de mando integral.

El componente de integración de datos (Data Integration) con la herramienta BusinessObjects Data Integrator v11.5 se utilizará para realizar los procesos de extracción, transformación y carga (ETL). Esta herramienta de integración de datos tiene como características fundamentales las siguientes:

- Conexión a cualquier base de datos a través de ODBC.
- Conexión a las bases de datos comerciales más importantes. Entre ellas Oracle, Microsoft SQL Server, y DB2.
- Carga de archivos planos separados por comas (CSV) y archivos XML.

- Herramienta gráfica para la creación de los flujos de carga. Permite arrastrar las tablas o archivos fuente, utilizar transformadores y enviar los resultados a las tablas destinos.
- Contiene transformadores que permiten aplicar la técnica SCD (Slowly Changing Dimension) Tipo 2 [8] que consiste en mantener un histórico de los cambios de atributos en las dimensiones agregando una columna que indica el estado del registro, sí está activo o no.
- Contiene un módulo Web de calendarización de flujos de carga a través del cuál se puede programar las fechas y horas en las cuales se cargarán los diversos flujos de carga construidos.

La entrega de información a través de la plataforma de inteligencia de negocios BOE XI R2 SP3 se la realizará a través del portal de BI BusinessObjects InfoView. Este portal Web constituye el centro de entrega de información a los usuarios finales. El portal mantiene los reportes de análisis en una estructura de directorios previamente creada, permitiendo asignar los permisos de acceso necesarios. Además, a través del portal se puede acceder a los tableros de mando integral y visualizar las diferentes métricas y analíticas creadas.

Los reportes y los tableros de mando integral utilizarán como fuente de datos el Data Mart para el ramo de vehículos que se construirá siguiendo las fases de desarrollo descritas en éste capítulo. Dentro de las herramientas de BOE XI R2 SP3, un Data Mart es conocido como Universo. Cada universo constituye un modelo de datos fuente para la elaboración de reportes o tableros. La herramienta de la plataforma que se usará para éste propósito se conoce como BusinessObjects Designer.

Los reportes con información para análisis del negocio serán construidos usando la herramienta BusinessObjects Web Intelligence y los tableros de mando integral serán construidos usando la herramienta BusinessObjects Performance Management.

[8] KIMBALL, Ralph. **The data warehouse toolkit: the complete guide to dimensional modeling.**

Se usará la base de datos Oracle 10g Release 2 para almacenar los datos fuente, los metadatos de las herramientas de la plataforma BOE XI R2 SP3 y el Data Mart para el ramo de vehículos.

En la Figura 3.1 se ilustran las herramientas de la plataforma BOE XI R2 SP3, las bases de datos que se utilizarán para los procesos de carga ETL, los metadatos y la base de datos en la cuál se almacenará el Data Mart.

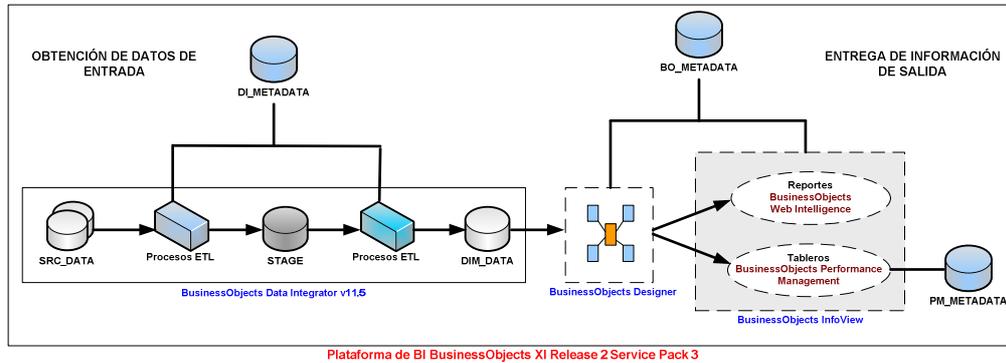


Figura 3.1 Herramientas de la plataforma BOE XI R2 SP3

Como se puede observar en la Figura 3.1, para la construcción del Data Mart se utilizará una arquitectura de flujo de datos simple con un almacén de datos dimensional (Simple DDS) descrita en el capítulo 1.2.4.1. Se extraerán los datos de la base de datos SRC_DATA a la base de datos STAGE. En la base de datos STAGE se realizarán las operaciones de transformación, limpieza y se aplicará la técnica SCD Tipo 2 para cargar el almacén de datos dimensional DIM_DATA. Los procesos ETL serán apoyados por la herramienta BusinessObjects Data Integrator v11.5, la misma que almacenará sus metadatos en la base de datos DI_METADATA.

Después de alimentar de datos al almacén de datos dimensional DIM_DATA, se construirá el universo de datos usando la herramienta BusinessObjects Designer. El universo de datos creado servirá como fuente para elaborar los reportes y tableros que permitan analizar el ramo de vehículos para una empresa aseguradora y tomar decisiones estratégicas, tácticas y operacionales.

Las herramientas BusinessObjects Designer y BusinessObjects Web Intelligence, almacenarán sus metadatos en la base de datos BO_METADATA. La herramienta BusinessObjects Performance Management almacenará sus metadatos en la base de datos PM_METADATA. Entiéndase como metadatos a datos necesarios que manejarán las herramientas de BI para su correcto funcionamiento.

3.1.2 RECURSOS

Las herramientas de la plataforma de inteligencia de negocios BOE XI R2 SP3 y el servidor de base de datos Oracle 10g serán instaladas en un computador portátil de las siguientes características:

Características de Hardware	
Característica	Descripción
Marca:	Acer
Modelo:	Aspire 5920
Procesador:	Intel® Core™ 2 Duo 1.83 [Ghz]
Memoria RAM:	3 [Gb]
Disco Duro:	232.88 [Gb]
Características de Software	
Característica	Descripción
Sistema Operativo:	Microsoft Windows XP Profesional Service Pack 3
Servidor de Base de datos:	Oracle Database 10g Release 2 (10.2) for Microsoft Windows (32-Bit)
Herramientas de Inteligencia de negocios:	BusinessObjects Data Integrator v 11.5
	BusinessObjects Enterprise XI Release 2 Service Pack 3
Herramienta de Modelamiento de datos:	AllFusion® ERwin® Data Modeler v7.1

Herramienta de Gestión de Oracle:	TOAD for Oracle® v9.0.1.8
Máquina virtual:	Java™ SE Development Kit 6 Update 7
Navegador Web:	Microsoft® Internet Explorer v6.0

Tabla 3.1 Características de Hardware y Software del equipo que se usará para la construcción del Data Mart para el ramo de vehículos

3.1.3 ACTIVIDADES A REALIZAR

Para la construcción del Data Mart se utilizará una metodología en cascada, realizando las siguientes actividades.

1. Preparar el ambiente de desarrollo del Data Mart, realizando las siguientes actividades:
 - Instalar el servidor de base de datos Oracle 10g R2.
 - Instalar la herramienta de gestión de Oracle, TOAD for Oracle v9.0.1.8.
 - Instalar la herramienta de modelamiento de datos ERwin Data Modeler v7.1.
 - Instalar el servidor de integración de datos BusinessObjects Data Integrator v11.5.
 - Instalar la plataforma de inteligencia de negocios BusinessObjects Enterprise XI Release 2 Service Pack 3 con la opción de servidor, incluyendo la herramienta de gestión del rendimiento (BusinessObjects Performance Management).

2. Definir requerimientos funcionales y no funcionales del Data Mart para el ramo de vehículos.

3. Analizar la base de datos fuente y comprender la estructura que será base para la construcción del Data Mart. Los entregables de ésta actividad serán:
 - Modelo lógico de las tablas que servirán como fuente para cargar el Data Mart, éste modelo se lo nombrará como ERD_SEGUROS_VEHICULOS.
 - Diccionario de datos del modelo ERD_SEGUROS_VEHICULOS, conteniendo claves primarias, claves foráneas, tablas con sus atributos y tipos de datos.

4. Diseñar el modelo dimensional del Data Mart en base al análisis de los requerimientos funcionales. Los entregables de ésta actividad serán:
 - Modelo dimensional de datos para el ramo de vehículos que abstraiga los requerimientos funcionales. Este modelo constituirá el modelo para el Data Mart de vehículos y se lo nombrará como DM_SEGUROS_VEHICULOS.
 - Diccionario de datos del modelo DM_SEGUROS_VEHICULOS.

5. Implementar los procesos para cargar el Data Mart DM_SEGUROS_VEHICULOS utilizando la herramienta de integración de datos BusinessObjects Data Integrator v11.5, considerando las siguientes actividades:
 - Preparar el esquema de Oracle SRC_DATA que contendrá los datos fuente para cargar el Data Mart.
 - Construir los procesos ETL para extraer, transformar, limpiar y cargar los datos en el Data Mart DM_SEGUROS_VEHICULOS, el cuál estará físicamente en el esquema de Oracle DIM_DATA.

6. Realizar las pruebas de carga evaluando la integridad referencial y que no exista fuga de datos en los procesos de carga para asegurar la confiabilidad de los datos del Data Mart DM_SEGUROS_VEHICULOS.

7. Crear el universo perteneciente al Data Mart DM_SEGUROS_VEHICULOS usando la herramienta BusinessObjects Designer.

8. Construir los reportes necesarios que cumplan con los requerimientos funcionales del Data Mart DM_SEGUROS_VEHICULOS. Además, construir el tablero de mando integral con los indicadores de gestión más relevantes, demostrando la funcionalidad del Data Mart y cómo puede apoyar a la toma de decisiones para el ramo de vehículos en una empresa aseguradora.

3.2 DEFINICIÓN DE REQUERIMIENTOS

3.2.1 REQUERIMIENTOS FUNCIONALES

Nro.	Requerimiento	Prioridad
1.	Los usuarios del negocio necesitan analizar los ingresos por primas a través del análisis de la prima suscrita, prima cobrada y prima pendiente de pago en cada día, mes y año por agencia, unidad de negocio, agente de seguros y cliente.	Alta
2.	Los usuarios del negocio necesitan analizar los egresos por concepto de siniestros a través del valor pagado en reclamos para cada día, mes y año por agencia, unidad de negocio, agente de seguros y cliente.	Alta
3.	Los usuarios del negocio necesitan analizar el resultado técnico para cada mes y año. Entiéndase como resultado técnico la suma de ingresos menos los egresos. Los ingresos constituyen las primas cobradas, y los egresos las comisiones pagadas al agente de seguros más los siniestros pagados y gastos administrativos.	Alta
4.	Los usuarios del negocio necesitan analizar la suma asegurada de la cartera para cada día, año y mes por agencia, unidad de negocio y agente de seguros.	Alta
5.	Los usuarios del negocio necesitan analizar el número de vehículos asegurados por agencia, unidad de negocio, agente de seguros, marca y modelo de vehículo.	Alta
6.	Los usuarios del negocio necesitan analizar el número de vehículos siniestrados por agencia, unidad de negocio, agente de seguros, marca y modelo de vehículo.	Alta
7.	Los usuarios del negocio necesitan analizar el índice de siniestralidad por agencia, unidad de negocio, agente de seguros, cliente, marca y modelo de vehículos.	Alta
8.	Los usuarios del negocio necesitan conocer la tasa promedio con la cuál se están asegurando los vehículos por agencia, unidad de negocio y agente. Entiéndase como tasa promedio, la división entre prima cobrada sobre valor asegurado.	Alta

9.	Los usuarios del negocio necesitan tener consolidado en un tablero de mando integral los indicadores de gestión para el ramo de vehículos más relevantes mostrándose las tendencias altas o bajas en un solo vistazo para cada agencia. Además, se requiere que puedan tener acceso al detalle de cada indicador para detectar causas por las bajas y tomar acciones al respecto.	Alta
10.	Los usuarios del negocio necesitan conocer el número de vehículos asegurados y siniestrados por cobertura. Además el índice de siniestralidad por cobertura en cada agencia y unidad de negocio.	Media
11.	Los usuarios del negocio necesitan conocer la cartera de vehículos asegurados y siniestrados. Conociendo para cada vehículo su marca, modelo, año de fabricación, placas, cilindraje, tonelaje, número de ocupantes, tipo de vehículo, color de vehículo, clase de vehículo y los extras asegurados.	Media
12.	Los usuarios del negocio necesitan analizar el valor asegurado promedio de la cartera para cada día, mes y año por agencia, unidad de negocio y agente de reaseguros. Además, les interesa analizar este indicador por cobertura.	Media
13.	Los usuarios del negocio requieren analizar las causas por las cuáles se han siniestrado los vehículos. Así como también el índice de intensidad de siniestros y la frecuencia para cada día, mes y año por agencia y unidad de negocio. Además se requiere realizar este análisis por marca y modelo de vehículo.	Media
14.	Los usuarios del negocio necesitan conocer el valor de reservas incrementadas y disminuidas por concepto de pago de siniestros para cada día, mes y año por agencia, unidad de negocio, agente de seguros y cliente.	Media
15.	Se requiere conocer la cartera de pólizas de vehículos asegurados obteniéndose el número de póliza, el nombre del cliente de la póliza, los asegurados, y beneficiarios. Además se requieren conocer los anexos o endosos de las pólizas y los estados por los cuáles han pasado.	Media

16.	Se requiere que el Data Mart almacene tres años de datos históricos para realizar las diversas actividades de análisis. Conteniendo datos desde el año 2005 hasta el año 2007.	Media
17.	Los usuarios del negocio necesitan conocer los porcentajes de comisiones respecto a las primas cobradas por día, mes y año para cada agencia, unidad de negocio y agente de seguros con el objeto de analizar las condiciones de negociación con los agentes de seguros.	Media
18.	Se requiere conocer el porcentaje de gastos administrativos con respecto a las primas cobradas por día, mes y año por agencia y unidad de negocio con el objeto de tomar decisiones, encontrar causas y disminuir los gastos administrativos y generar más utilidad al negocio.	Media
19.	Se requiere conocer la tasa de renovación de pólizas por agencia, unidad de negocio, agente de seguros y cliente.	Baja
20.	Se requiere analizar el tiempo que toma en tramitarse un siniestro. Considerando la fecha en la que ocurrió, la fecha en la que se abrieron las reservas y la fecha en la que se pagó la indemnización con el objeto de mejorar el servicio de tramitación de siniestros.	Baja
21.	Se requieren reportes que puedan ser visualizados a través de dispositivos móviles y que contengan los indicadores de gestión más relevantes como son primas cobradas, pagos de siniestros, siniestralidad, número de vehículos asegurados y siniestrados para cada agencia y unidad de negocio.	Baja

Tabla 3.2 Tabla de requerimientos funcionales del Data Mart

3.2.2 REQUERIMIENTOS NO FUNCIONALES

Nro.	Requerimiento
1.	Los reportes de análisis y el tablero de mando integral deberán tener una interfaz fácil de usar y basada en la Web.
2.	Los usuarios deben tener la posibilidad de mostrar gráficos o tablas en los reportes, así como también las tendencias de los mismos. Además, se requiere que los reportes puedan ser impresos o en su defecto que puedan ser exportados a archivos de Excel o PDF.
3.	El portal de inteligencia de negocios deberá estar basado en la Web, y sólo podrá ser accedido a través de la red interna de la compañía. No se podrá tener acceso a través del Internet a pesar de que sea basada en la Web.
4.	Las herramientas de reporte deben permitir a los usuarios realizar consultas ad hoc sin necesidad de escribir código SQL. Los usuarios deben tener la posibilidad de crear reportes a demanda de una forma fácil y rápida.
5.	En los reportes los usuarios deben tener la posibilidad de realizar agregaciones y desagregaciones a través de jerarquías de datos. Es decir los usuarios podrán realizar un análisis desde lo general hasta lo específico o viceversa.
6.	En los reportes los usuarios deberán tener la posibilidad de tener alertas, crear bloques seccionados, y realizar un TOP N de las mejores o peores agencias, unidades de negocios, agentes de seguros, clientes, marcas y vehículos analizando su siniestralidad.
7.	Las herramientas de inteligencia de negocios deben permitir la fácil recuperación ante fallos y ser lo suficientemente flexibles para sacar respaldos y poder ser recuperadas sin problemas.

Tabla 3.3 Tabla de requerimientos no funcionales del Data Mart

3.3 DISEÑO

En ésta fase se diseñará el modelo dimensional para el Data Mart para el ramo de vehículos. Primero, es necesario comprender la estructura de la base de datos que servirá como fuente para cargar el Data Mart. Se describirá la base de datos fuente de seguros y su diccionario de datos; explicándose, tablas, atributos y relaciones. Finalmente, se realizará el diseño dimensional del Data Mart para el ramo de vehículos en base al análisis de los requerimientos funcionales.

3.3.1 BASE DE DATOS DE SEGUROS DE VEHÍCULOS

3.3.1.1 Modelo de datos lógico

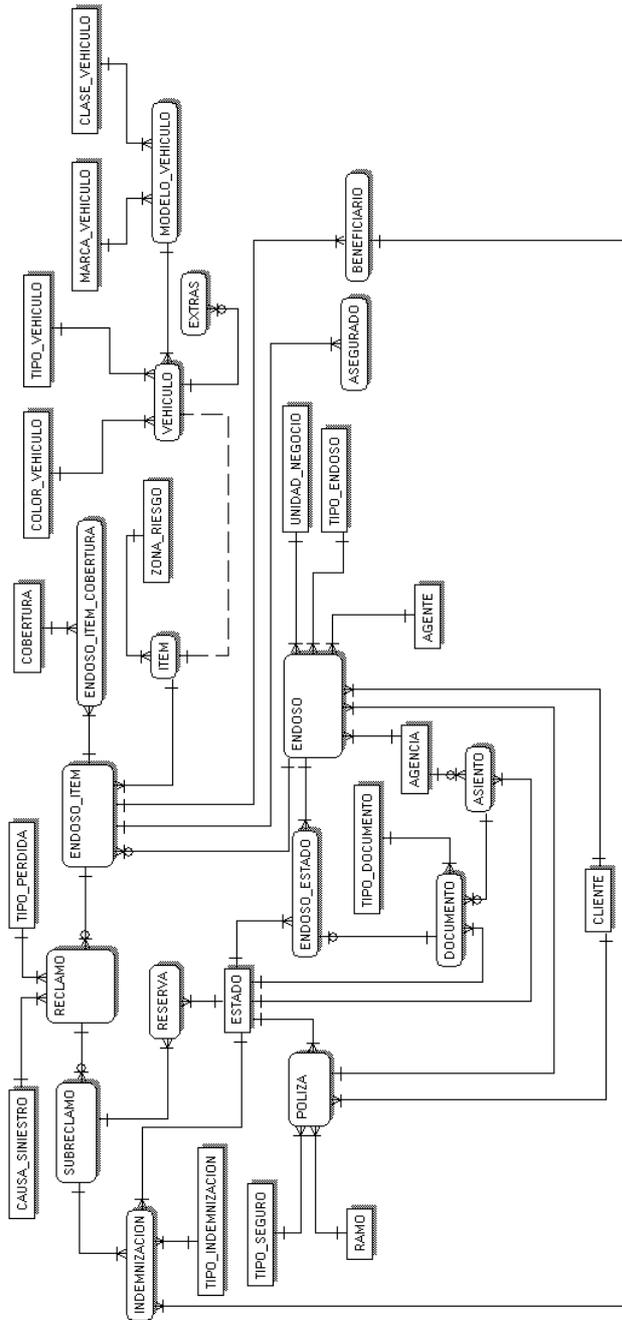


Figura 3.2 Modelo de datos lógico para seguros de vehículos

En la Figura 3.2 se ilustra el modelo de datos lógico para seguros de vehículos. Los seguros de vehículos se formalizan a través de una póliza. Una póliza

pertenece a un ramo, en éste caso, el ramo de vehículos. La póliza puede ser de un tipo de seguro sea este normal, coaseguro o facultativo. La póliza tiene un cliente y puede tener uno o más anexos. Los anexos de la póliza son conocidos como endosos. Un endoso es un accesorio de la póliza y puede pasar por varios estados que van desde la emisión hasta el pago del mismo. El endoso pertenece a un cliente y es emitido en una unidad de negocio de una agencia con o sin participación de un agente de seguros. El endoso puede ser de un tipo sea éste una inclusión, exclusión, aumento de suma asegurada o prima, etc.

Un endoso nace en estado de “Borrador”, en la emisión pasa a un estado “Emitido” y a un estado de “Facturado” cuando se ha generado la factura. La factura constituye un documento financiero. Después de que el endoso ha sido pagado, éste pasa al estado de “Pagado” y se genera un asiento contable con el valor del mismo. Sí el endoso es cancelado o anulado pasa a un estado de “Cancelado” generándose una nota de crédito. Los documentos pueden ser de tipo factura o nota de crédito.

Un endoso puede tener uno o más objetos asegurados o ítems. El ítem, para éste caso de estudio es representado por un vehículo. Además, cada ítem tiene una zona de riesgo que corresponde a los límites territoriales en los cuáles está asegurado un vehículo. El vehículo puede tener o no asegurado extras como por ejemplo llanta de emergencia o radios. El vehículo tiene color, es de un tipo sea ésta camioneta, automóvil, bus, jeep etc. Además, el vehículo tiene una marca, modelo y puede pertenecer a una clase sea ésta motos, vehículos livianos o vehículos pesados.

Cada endoso y sus objetos asegurados tienen una o varias coberturas. La cobertura representa bajo qué riesgo está cubierto un vehículo, éstas pueden ser las coberturas de robo total o parcial, choque total o parcial, todo riesgo, choque y volcadura, etc. Además, cada endoso y su objeto asegurado tienen un asegurado y un beneficiario del seguro. Sí el vehículo asegurado de un endoso sufre un siniestro se genera un reclamo, el mismo que tiene una causa y un tipo de pérdida. Por ejemplo la causa de un vehículo siniestrado podría ser la volcadura y tener un tipo de pérdida total por daño. El reclamo genera uno o varios subreclamos por cada cobertura que se indemnizará. Cada subreclamo genera

una o varias reservas que corresponden a una cantidad de dinero que servirá para afrontar la indemnización. La indemnización se genera cuando cada subreclamo ha sido pagado al beneficiario respectivo. La póliza, el endoso, el documento, el asiento, la reserva y la indemnización tienen un estado.

3.3.1.2 Diccionario de datos

3.3.1.2.1 Tabla AGENCIA

Columna	Tipo de Dato	Descripción
ID	VARCHAR2(12)	Clave primaria de la tabla.
NOMBRE	VARCHAR2(40)	Nombre de la agencia.
ESPRINCIPAL	VARCHAR2(1)	Bandera con valores 1 o 0 que indica si la agencia es principal o no. 1 es Verdadero, 0 es falso.

Tabla 3.4 Tabla AGENCIA

En el Anexo G se continúa con el diccionario de datos de la base de datos de seguros de vehículos.

3.3.2 DATA MART PARA ANÁLISIS DE SEGUROS DE VEHÍCULOS

3.3.2.1 Modelo dimensional de datos

El diseño del modelo dimensional de datos para el Data Mart de seguros de vehículos se lo realizará enfocándose en las dimensiones e indicadores de análisis de producción de pólizas de vehículos y sus siniestros.

Para el diseño, se han utilizado las técnicas de modelamiento dimensional propuestas por Kimball. Se ha aplicado el concepto de dimensiones degeneradas (DD, Degenerate Dimension)[8] para el caso de las pólizas, reclamos y subreclamos. Una dimensión degenerada es aquella dimensión que únicamente tiene un atributo y alta granularidad. Éste atributo se lo incluye en la tabla de hechos. Además, se usarán claves subrogadas (SK, Subrogate Keys) para cargar las dimensiones. Las claves subrogadas son números enteros generados al

[8] KIMBALL, Ralph. *The data warehouse toolkit: the complete guide to dimensional modeling*.

aplicar la técnica SCD, éstas claves ayudan a mejorar el rendimiento en las juntas entre las dimensiones y las tablas de hechos.

En la Figura 3.3 se ilustra el modelo de datos dimensional diseñado para el análisis de producción de pólizas y siniestros para el ramo de vehículos en una empresa aseguradora.

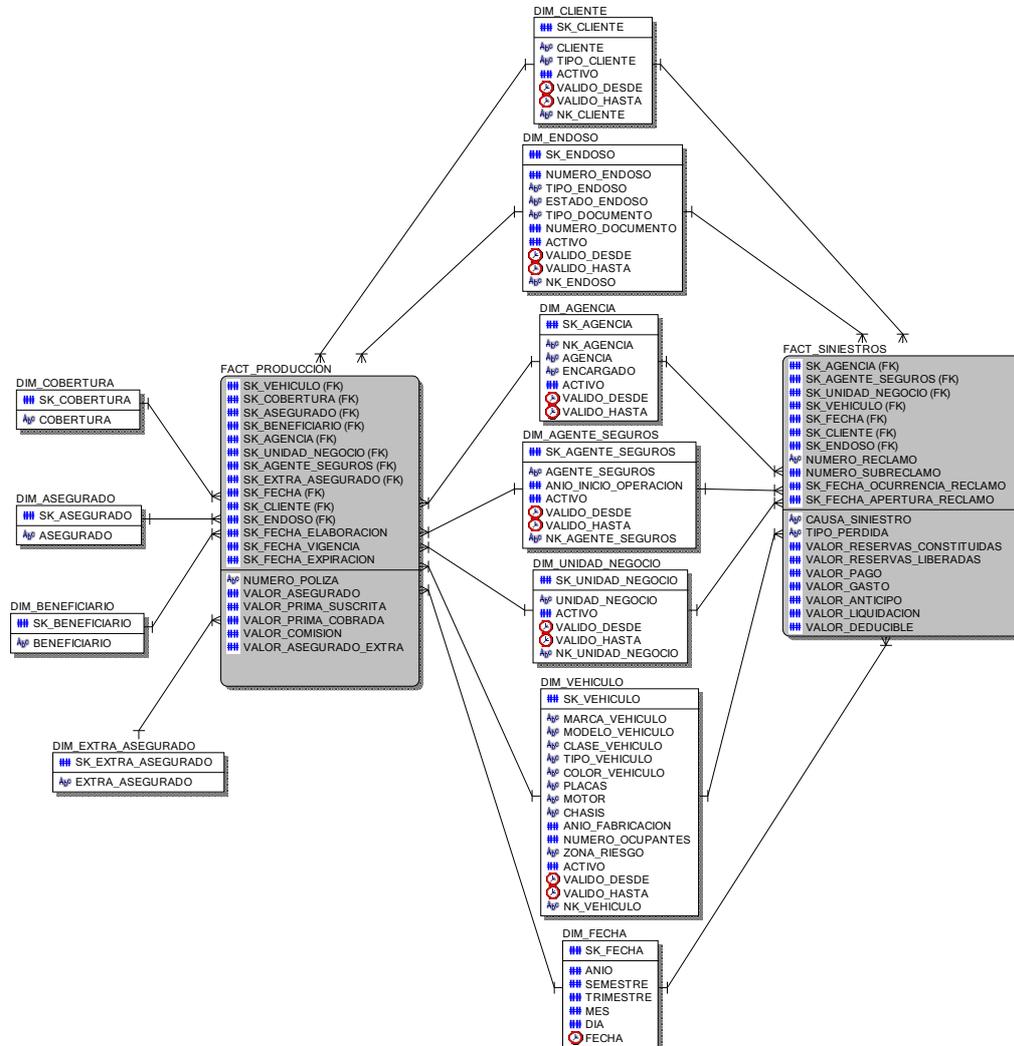


Figura 3.3 Modelo de datos dimensional para el Data Mart

Se ha utilizado un esquema de tablas de hechos tipo constelación (Fact Constellation Schema). Se creó dos tablas de hechos distintas con dimensiones compartidas. La primera tabla de hechos FACT_PRODUCION está relacionada

con las dimensiones de análisis de producción de pólizas. Ésta tabla posee como dimensión degenerada la póliza ya que posee un único atributo, el número de póliza.

La tabla de hechos FACT_SINIESTROS contiene los indicadores de análisis para siniestros por endoso de cada póliza. Esta tabla de hechos está relacionada con las dimensiones compartidas de producción de pólizas y contiene las dimensiones degeneradas que son el reclamo, subreclamo, causa del siniestro y tipo de pérdida. Se aplicó el mismo concepto para crear éstas dimensiones como degeneradas.

Tanto las tablas de hechos FACT_PRODUCCION y FACT_SINIESTROS son tablas transaccionales que contendrán cada registro de producción y siniestros por cada fecha de transacción. Las tablas de hechos permitirán realizar el análisis de ingresos y egresos para cada día, mes y año por agencia, unidad de negocio, agente de seguros y cliente cumpliéndose los requerimientos número 1, 2 y 3.

El último nivel de detalle por cada tabla de hechos es el endoso y su objeto asegurado, en nuestro caso de estudio el vehículo. El modelo dimensional permitirá realizar un análisis de producción y siniestros hasta el último nivel de detalle que será el vehículo que pertenece a un endoso de una póliza.

Cada registro de la tabla de hechos FACT_PRODUCCION corresponde a cada vehículo asegurado y la tabla de hechos FACT_SINIESTROS corresponde a cada vehículo siniestrado. Las dimensiones de análisis propias para la tabla de hechos FACT_PRODUCCION son la dimensión DIM_COBERTURA, DIM_ASEGURADO, DIM_BENEFICIARIO, DIM_EXTRA_ASEGURADO. Las dimensiones DIM_CLIENTE, DIM_ENDOSO, DIM_AGENCIA, DIM_AGENTE_SEGUROS, DIM_UNIDAD_NEGOCIO y DIM_FECHA son conocidas como dimensiones conformadas (Conformed Dimensions) [8] ya que pueden ser compartidas entre las diversas tablas de hechos creadas para un mismo Data Mart o para otros, siempre y cuando se compartan los mismos datos. En el siguiente subcapítulo se

[8] KIMBALL, Ralph. *The data warehouse toolkit: the complete guide to dimensional modeling*.

describe el diccionario de datos para el modelo dimensional del Data Mart diseñado.

3.3.2.2 Diccionario de datos

3.3.2.2.1 Dimensión DIM_AGENCIA

La dimensión agencia representa la agencia en la cuál se suscribieron las pólizas. Esta dimensión permitirá realizar un análisis de producción y siniestros por agencia para evaluar sí es que la agencia está suscribiendo riesgos que le resulten rentables o no para la compañía aseguradora.

Columna	Tipo de Dato	Descripción
SK_AGENCIA	INTEGER	Clave subrogada de la dimensión. Esta clave se generará en los procesos ETL después de aplicar la técnica de control de cambios SCD Tipo 2.
AGENCIA	VARCHAR2(40)	Nombre de la agencia.
ENCARGADO	VARCHAR2(80)	Gerente responsable de la Agencia, responsable de la gestión y manejo de la agencia.
ACTIVO	INTEGER	Bandera para indicar sí el registro es válido o sufrió algún cambio después de aplicar la técnica SCD Tipo 2. 1 sí el registro está activo o 0 sí no lo está.
VALIDO_DESDE	DATE	Fecha desde la cual es válido el registro.
VALIDO_HASTA	DATE	Fecha hasta la cual es válido el registro.
NK_AGENCIA	VARCHAR2(12)	Clave natural de la agencia. Es preservada para realizar el control de cambios en base a ésta clave.

Tabla 3.5 Dimensión DIM_AGENCIA

3.3.2.2.2 Dimensión DIM_AGENTE_SEGUROS

La dimensión agente de seguros representa la empresa intermediaria que trae clientes a las agencias para la suscripción de pólizas. El análisis por ésta dimensión permitirá evaluar qué tan rentables o no resultan los clientes traídos, para de ésta manera aumentar sus comisiones o disminuirlas.

Columna	Tipo de Dato	Descripción
SK_AGENTE_SEGUROS	INTEGER	Clave subrogada de la dimensión.
AGENTE_SEGUROS	VARCHAR2(120)	Nombre del intermediario de seguros.
ANIO_INICIO_OPERACION	INTEGER	Año desde el cuál el agente empezó a operar legalmente.
ACTIVO	INTEGER	Bandera para indicar sí el registro es válido o sufrió algún cambio después de aplicar la técnica SCD Tipo 2. 1 sí el registro está activo o 0 sí no lo está.
VALIDO_DESDE	DATE	Fecha desde la cual es válido el registro.
VALIDO_HASTA	DATE	Fecha hasta la cual es válido el registro.
NK_AGENTE_SEGUROS	VARCHAR2(12)	Clave natural de la tabla.

Tabla 3.6 Dimensión DIM_AGENTE_SEGUROS

3.3.2.2.3 Dimensión DIM_ASEGURADO

La dimensión asegurado representa el nombre del asegurado de un vehículo asegurado en un endoso.

Columna	Tipo de Dato	Descripción
SK_ASEGURADO	INTEGER	Clave subrogada de la dimensión.
ASEGURADO	VARCHAR2(120)	Nombre del asegurado.

Tabla 3.7 Dimensión DIM_ASEGURADO

3.3.2.2.4 Dimensión DIM_BENEFICIARIO

La dimensión representa el nombre del o los beneficiarios a los cuáles se les pagará una indemnización sí es que un vehículo asegurado en un endoso se siniestra.

Columna	Tipo de Dato	Descripción
SK_BENEFICIARIO	INTEGER	Clave subrogada de la dimensión.
BENEFICIARIO	VARCHAR2(120)	Nombre del beneficiario.

Tabla 3.8 Dimensión DIM_BENEFICIARIO

3.3.2.2.5 Dimensión DIM_CLIENTE

La dimensión cliente representa al cliente que tomó el seguro. Esta dimensión permitirá analizar el resultado técnico de la compañía de seguros por cliente, de tal forma que se puedan conocer los mejores y los peores clientes para aplicar beneficios o castigos en las tasas de suscripción de nuevas pólizas.

Columna	Tipo de Dato	Descripción
SK_CLIENTE	INTEGER	Clave subrogada de la dimensión.
CLIENTE	VARCHAR2(120)	Nombre del cliente.
TIPO_CLIENTE	VARCHAR2(15)	Representa el tipo del cliente. Indica sí el cliente es una empresa o es una persona natural.
ACTIVO	INTEGER	Bandera para indicar sí el registro es válido o sufrió algún cambio después de aplicar la técnica SCD Tipo 2. 1 sí el registro está activo o 0 sí no lo está.
VALIDO_DESDE	DATE	Fecha desde la cual es válido el registro.
VALIDO_HASTA	DATE	Fecha hasta la cual es válido el registro.
NK_CLIENTE	VARCHAR2(12)	Clave natural de la tabla.

Tabla 3.9 Dimensión DIM_CLIENTE

3.3.2.2.6 Dimensión DIM_COBERTURA

La dimensión cobertura representa las coberturas por la cuáles está asegurado un vehículo en un endoso.

Columna	Tipo de Dato	Descripción
SK_COBERTURA	INTEGER	Clave subrogada de la dimensión.
COBERTURA	VARCHAR2(40)	Nombre de la cobertura.

Tabla 3.10 Dimensión DIM_COBERTURA

3.3.2.2.7 Dimensión DIM_ENDOSO

La dimensión endoso representa el anexo por póliza creado, el mismo que contiene uno o más vehículos asegurados. El endoso tiene asociado un documento que puede ser una factura sí el endoso está vigente o una nota de crédito sí el endoso ha sido cancelado.

Columna	Tipo de Dato	Descripción
SK_ENDOSO	INTEGER	Clave subrogada de la dimensión.
NUMERO_ENDOSO	NUMBER(10)	Número del endoso o anexo de la póliza.
TIPO_ENDOSO	VARCHAR2(40)	Tipo de endoso. Indica sí el endoso es una exclusión, inclusión, cancelación, etc.
ESTADO_ENDOSO	VARCHAR2(20)	Indica el estado del endoso.
TIPO_DOCUMENTO	VARCHAR2(20)	Indica el tipo de documento que está asociado al endoso. Puede ser una factura o una nota de crédito.
NUMERO_DOCUMENTO	NUMBER(10)	Indica el número de documento asociado.
ACTIVO	INTEGER	Bandera para indicar sí el registro es válido o sufrió algún cambio después de aplicar la técnica SCD Tipo 2. 1 sí el registro está

		activo o 0 sí no lo está.
VALIDO_DESDE	DATE	Fecha desde la cual es válido el registro.
VALIDO_HASTA	DATE	Fecha hasta la cual es válido el registro.
NK_ENDOSO	VARCHAR2(12)	Clave natural de la tabla.

Tabla 3.11 Dimensión DIM_ENDOSO

3.3.2.2.8 Dimensión DIM_EXTRA_ASEGURADO

La dimensión extra asegurado representan los extras asegurados por cada vehículo. Los extras pueden ser el blindaje del vehículo, radio, reproductor de video o las llantas de emergencia.

Columna	Tipo de Dato	Descripción
SK_EXTRA_ASEGURADO	INTEGER	Clave subrogada de la dimensión.
EXTRA_ASEGURADO	VARCHAR2(120)	Nombre del extra que se ha asegurado.

Tabla 3.12 Dimensión DIM_EXTRA_ASEGURADO

3.3.2.2.9 Dimensión DIM_FECHA

La dimensión fecha es una de las dimensiones que siempre está presente en un Data Mart. Esta dimensión permitirá realizar el análisis histórico de la producción de pólizas y los siniestros. La dimensión fecha contiene los datos de fecha en los cuáles se realizaron transacciones por producción o siniestros.

Columna	Tipo de Dato	Descripción
SK_FECHA	INTEGER	Clave subrogada de la dimensión.
ANIO	INTEGER	Año al cuál pertenece la fecha.
SEMESTRE	INTEGER	Número de semestre al cuál pertenece la fecha.
TRIMESTRE	INTEGER	Número de trimestre al cuál pertenece la fecha.
MES	INTEGER	Número de mes al cuál pertenece la

		fecha.
DIA	INTEGER	Día del mes al cuál pertenece la fecha.
FECHA	DATE	Fecha en la cuál se realizó una transacción de producción o siniestros.

Tabla 3.13 Dimensión DIM_FECHA

3.3.2.2.10 Dimensión DIM_UNIDAD_NEGOCIO

La dimensión unidad de negocio representa la unidad manejada por cada agencia para la suscripción de pólizas por negocios masivos o por negocios individuales.

Columna	Tipo de Dato	Descripción
SK_UNIDAD_NEGOCIO	INTEGER	Clave subrogada de la dimensión.
UNIDAD_NEGOCIO	VARCHAR2(40)	Nombre de la unidad de negocio manejada por cada agencia.
ACTIVO	INTEGER	Bandera para indicar sí el registro es válido o sufrió algún cambio después de aplicar la técnica SCD Tipo 2. 1 sí el registro está activo o 0 sí no lo está.
VALIDO_DESDE	DATE	Fecha desde la cual es válido el registro.
VALIDO_HASTA	DATE	Fecha hasta la cual es válido el registro.
NK_UNIDAD_NEGOCIO	VARCHAR2(12)	Clave natural de la tabla.

Tabla 3.14 Dimensión DIM_UNIDAD_NEGOCIO

3.3.2.2.11 Dimensión DIM_VEHICULO

La dimensión vehículo representa el objeto asegurado en un endoso de una póliza. Esta dimensión constituye el objeto de análisis principal para el caso de estudio. Permitirá realizar un análisis de producción y siniestros para cada uno de sus detalles de tal forma que apoye en la toma de decisiones necesarias que generen valor al negocio.

Columna	Tipo de Dato	Descripción
SK_VEHICULO	INTEGER	Clave subrogada de la dimensión.
COLOR_VEHICULO	VARCHAR2(40)	Color del vehículo.
MARCA_VEHICULO	VARCHAR2(20)	Marca del vehículo.
MODELO_VEHICULO	VARCHAR2(40)	Modelo del vehículo.
CLASE_VEHICULO	VARCHAR2(10)	Indica sí el vehículo es un vehículo liviano, pesado o una moto.
TIPO_VEHICULO	VARCHAR2(20)	Indica el tipo de vehículo, sí es un jeep, automóvil, camioneta, bus, etc.
ZONA_RIESGO	VARCHAR2(10)	Zona en la cuál está cubierto el vehículo.
PLACAS	VARCHAR2(10)	Placas del vehículo.
MOTOR	VARCHAR2(40)	Número de motor del vehículo.
CHASIS	VARCHAR2(40)	Número de chasis del vehículo.
ANIO_FABRICACION	INTEGER	Año en el cuál fue fabricado el vehículo.
NUMERO_OCUPANTES	INTEGER	Número de ocupantes del vehículo.
ACTIVO	INTEGER	Bandera para indicar sí el registro es válido o sufrió algún cambio después de aplicar la técnica SCD Tipo 2. 1 sí el registro está activo o 0 sí no lo está.
VALIDO_DESDE	DATE	Fecha desde la cual es válido el registro.
VALIDO_HASTA	DATE	Fecha hasta la cual es válido el registro.
NK_VEHICULO	VARCHAR2(12)	Clave natural de la tabla.

Tabla 3.15 Dimensión DIM_VEHICULO

3.3.2.2.12 Tabla de Hechos FACT_PRODUCION

Columna	Tipo de Dato	Descripción
SK_VEHICULO	INTEGER	Clave foránea que hace referencia a la dimensión DIM_VEHICULO.
SK_COBERTURA	INTEGER	Clave foránea que hacer referencia a la dimensión DIM_COBERTURA.
SK_ASEGURADO	INTEGER	Clave foránea que hace referencia a la dimensión DIM_ASEGURADO.
SK_BENEFICIARIO	INTEGER	Clave foránea que hace referencia a la dimensión DIM_BENEFICIARIO.
SK_AGENCIA	INTEGER	Clave foránea que hace referencia a la dimensión DIM_AGENCIA.
SK_UNIDAD_NEGOCIO	INTEGER	Clave foránea que hace referencia a la dimensión DIM_UNIDAD_NEGOCIO.
SK_AGENTE_SEGUROS	INTEGER	Clave foránea que hace referencia a al dimensión DIM_AGENTE_SEGUROS.
SK_EXTRA_ASEGURADO	INTEGER	Clave foránea que hace referencia a la dimensión DIM_EXTRA_ASEGURADO.
SK_FECHA	INTEGER	Clave foránea que hace referencia a la dimensión DIM_FECHA.
SK_CLIENTE	INTEGER	Clave foránea que hace referencia a la dimensión DIM_CLIENTE.
SK_ENDOSO	INTEGER	Clave foránea que hace referencia a la dimensión

		DIM_ENDOSO.
SK_FECHA_ELABORACION	INTEGER	Clave foránea que hace referencia a la dimensión DIM_FECHA.
SK_FECHA_VIGENCIA	INTEGER	Clave foránea que hace referencia a la dimensión DIM_FECHA.
SK_FECHA_EXPIRACION	INTEGER	Clave foránea que hace referencia a la dimensión DIM_FECHA.
NUMERO_POLIZA	VARCHAR2(12)	Número de la póliza.
VALOR_PRIMA_SUSCRITA	NUMBER(8,2)	Valor que se suscribió por endoso.
VALOR_PRIMA_COBRADA	NUMBER(8,2)	Valor que se ha pagado por endoso.
VALOR_ASEGURADO	NUMBER(8,2)	Valor asegurado por endoso.
VALOR_ASEGURADO_EXTRA	NUMBER(8,2)	Valor asegurado del extra sí es que el vehículo lo tiene.
VALOR_COMISION	NUMBER(8,2)	Valor que se ha pagado al agente de seguros por cada endoso.

Tabla 3.16 Tabla de Hechos FACT_PRODUCION

3.3.2.2.13 Tabla de Hechos FACT_SINIESTROS

Columna	Tipo de Dato	Descripción
SK_AGENCIA	INTEGER	Clave foránea que hace referencia a la dimensión DIM_AGENCIA.
SK_AGENTE_SEGUROS	INTEGER	Clave foránea que hace referencia a la dimensión DIM_AGENTE_SEGUROS.
SK_UNIDAD_NEGOCIO	INTEGER	Clave foránea que hace referencia a la dimensión DIM_UNIDAD_NEGOCIO.

SK_VEHICULO	INTEGER	Clave foránea que hace referencia a la dimensión DIM_VEHICULO.
SK_FECHA	INTEGER	Clave foránea que hace referencia a la dimensión DIM_FECHA.
SK_CLIENTE	INTEGER	Clave foránea que hace referencia a la dimensión DIM_CLIENTE.
SK_ENDOSO	INTEGER	Clave foránea que hace referencia a la dimensión DIM_ENDOSO.
SK_FECHA_OCURRENCIA_RECLAMO	INTEGER	Clave foránea que hace referencia a la dimensión DIM_FECHA.
SK_FECHA_APERTURA_RECLAMO	INTEGER	Clave foránea que hace referencia a la dimensión DIM_FECHA.
NUMERO_RECLAMO	VARCHAR2(50)	Número del reclamo.
NUMERO_SUBRECLAMO	INTEGER	Número del subreclamo.
TIPO_PERDIDA	VARCHAR2(50)	Tipo de pérdida que ocasionó el siniestro. Esta puede ser total, parcial, etc.
CAUSA_SINIESTRO	VARCHAR2(50)	Causa por la cuál se originó el siniestro.
VALOR_RESERVAS_CONSTITUIDAS	NUMBER(8,2)	Valor de las reservas que se constituyeron para afrontar el siniestro.
VALOR_RESERVAS_LIBERADAS	NUMBER(8,2)	Valor de las reservas que se liberaron después del pago del siniestro.
VALOR_GASTO	NUMBER(8,2)	Valor generado por gastos operativos ocasionados en el pago del siniestro.

VALOR_DEDUCIBLE	NUMBER(8,2)	Valor del deducible que se cobró al cliente por subreclamo.
VALOR_PAGO	NUMBER(8,2)	Valor de pago del siniestro.
VALOR_LIQUIDACION	NUMBER(8,2)	Valor de liquidación del siniestro.
VALOR_ANTICIPO	NUMBER(8,2)	Valor pagado como anticipo del siniestro.

Tabla 3.17 Tabla de Hechos FACT_SINIESTROS

3.4 CONSTRUCCIÓN DEL DATA MART

En éste subcapítulo se explican los procesos de carga de datos del Data Mart DM_SEGUROS_VEHICULOS diseñado. La construcción de los flujos de carga del Data Mart serán realizados usando la herramienta de integración de datos BusinessObjects Data Integrator v11.5.

3.4.1 ARQUITECTURA DE FLUJO DE DATOS

De acuerdo a lo mencionado en los subcapítulos anteriores, se usará una arquitectura de flujo de datos simple usando la herramienta BusinessObjects Data Integrator v11.5 como se muestra en la Figura 3.4.

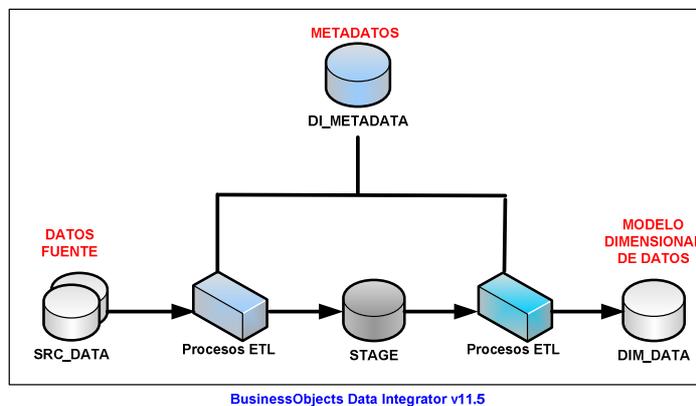


Figura 3.4 Arquitectura de flujo de carga de datos que se implementará en la herramienta BusinessObjects Data Integrator.

En la base de datos SRC_DATA se encuentra la base de datos fuente de seguros de vehículos descrita en el capítulo 3.3.1. Los procesos de extracción, transformación y carga de datos (ETL) serán implementados usando los transformadores de datos y demás objetos que la herramienta posee. Se extraerán los datos de la base de datos SRC_DATA a la base de datos STAGE. La base de datos STAGE servirá como base de datos temporal para realizar las operaciones de transformación de datos, limpieza, creación de claves subrogadas y creación de tablas temporales previas a la carga de datos del Data Mart que se encontrará en la base de datos DIM_DATA.

La definición de los procesos de carga y los metadatos manejados por la herramienta BusinessObjects Data Integrator se almacenarán en la base de datos DI_METADATA.

3.4.2 HERRAMIENTA BUSINESSOBJECTS DATA INTEGRATOR

3.4.2.1 Introducción

BusinessObjects Data Integrator utiliza un repositorio en dónde se almacenan todos los metadatos de la herramienta. En éste proyecto se usará el esquema de Oracle DI_METADATA. El acceso a la herramienta dependerá de la contraseña creada para el repositorio o esquema de Oracle, en la Figura 3.5 se muestra la pantalla de acceso a la herramienta.

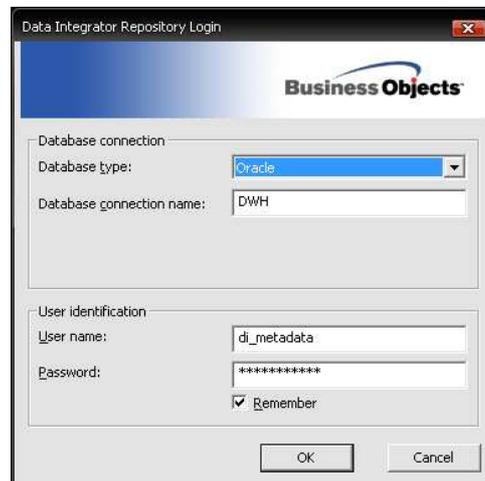


Figura 3.5 Pantalla de acceso a Data Integrator

Data Integrator contiene varios componentes que apoyan a los procesos de extracción, transformación y carga de un Data Warehouse. Uno de estos componentes es el Diseñador (Data Integrator Designer), el diseñador es una herramienta gráfica para la creación de flujos de carga de datos. La pantalla principal del diseñador contiene un área de proyectos, una barra de menú, una barra de herramientas, una paleta de herramientas, un área de trabajo y una librería de objetos locales, en la Figura 3.6 se muestra la pantalla principal de Data Integrator Designer.

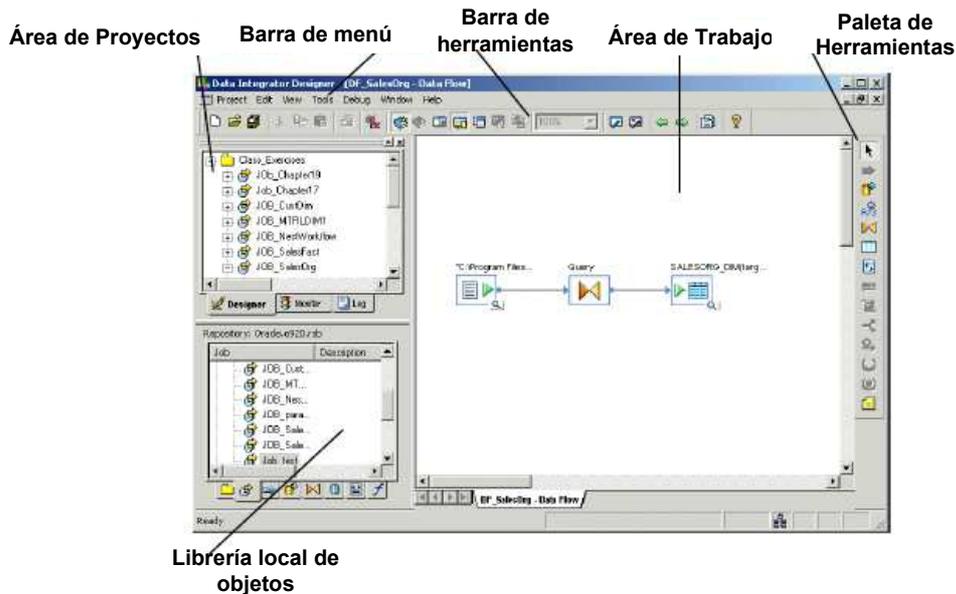


Figura 3.6 Pantalla Principal de Data Integrator Designer

El diseñador, Data Integrator Designer, maneja una jerarquía de objetos para gestionar los proyectos y sus flujos de carga de datos, como se muestra en la Figura 3.7.

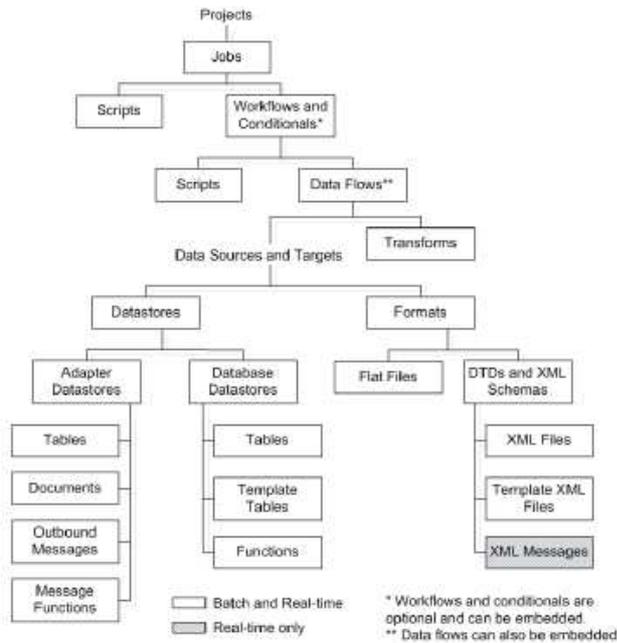


Figura 3.7 Jerarquía de objetos manejados por Data Integrator Designer

En Data Integrator los procesos ETL son gestionados por proyectos (Projects). Un proyecto puede tener varios Jobs. El Job es el único objeto que puede ejecutarse, es el punto de ejecución de un flujo de carga de datos. Un Job puede ser ejecutado en la herramienta de diseño para realizar pruebas y en producción para calendarizar su ejecución.

Los Jobs pueden contener scripts de carga o flujos de trabajo (Work Flows). Un flujo de trabajo define el proceso de toma de decisiones para ejecutar flujos de datos (Data Flows) determinando su camino de ejecución. Los flujos de trabajo pueden contener scripts o varios flujos de datos (Data Flows). Un flujo de datos extrae, transforma y carga los datos.

Los flujos de datos (Data Flows) contienen transformadores. Un transformador es un objeto propio de Data Integrator con funciones específicas como por ejemplo generar una clave, comparar tablas, controlar cambios, hacer uniones, etc. Los flujos de datos extraen datos de las fuentes de datos (Data Sources) y los cargan en los destinos (Targets).

Las fuentes y los destinos (Data Sources and Targets) pueden ser accedidos a través de almacenes de datos (Datastores) o formatos (Formats). Los almacenes de datos de base de datos (Database datastores) contienen las tablas, tablas temporales y funciones que se usarán para extraer y cargar los datos tanto en las fuentes como en los destinos. Los formatos (Formats) permiten cargar archivos planos y archivos XML.

3.4.2.2 Arquitectura de Data Integrator

La herramienta de integración de datos Business Objects Data Integrator contiene componentes estándar, componentes opcionales y herramientas de administración. Los componentes estándar de Data Integrator se ilustran en la Figura 3.8.

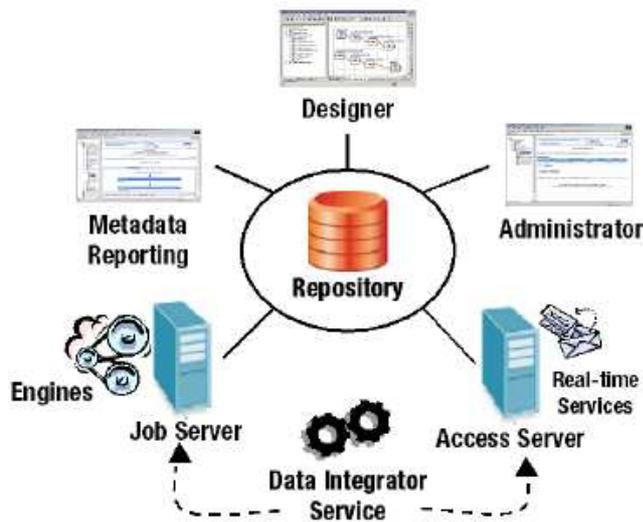


Figura 3.8 Componentes Estándar de Data Integrator [35]

El diseñador (Designer) es una herramienta de desarrollo que permite definir aplicaciones de gestión de datos que consisten de mapeo de datos, transformaciones y lógica de control [35]. Estas aplicaciones son representadas por proyectos del diseñador que contienen flujos de trabajo (Work Flows) y flujos de datos (Data Flows). Los objetos del diseñador representan los metadatos almacenados en el repositorio DI_METADATA.

[35] Business Objects. **Data Integrator Getting Started Guide**.

El repositorio de Data Integrator, DI_METADATA para el proyecto, es un conjunto de tablas que contiene los objetos creados por el diseñador y los objetos predefinidos del sistema, los metadatos de las fuentes y destinos, y las reglas de transformación. Cada repositorio está asociado con uno o más Data Integrator Job Servers, los cuáles ejecutan los Jobs que se crean a través del diseñador.

El Data Integrator Job Server inicia el motor (Engine) de movimiento de datos que integra datos de múltiples fuentes, realiza transformaciones complejas de datos, y carga los datos en los diferentes destinos definidos. El movimiento de datos puede ser a través de procesos por lotes (batch) o en tiempo real. En un ambiente de producción, el Job Server ejecuta Jobs disparados por un planificador o por un servicio de tiempo real manejado por el Data Integrator Access Server.

El administrador de Data Integrator (Data Integrator Administrador) provee de tareas de administración del sistema vía Web. Entre las tareas de administración se incluye, la planificación, monitoreo y ejecución de trabajos por lotes (batch jobs), gestión de servicios de tiempo real y administración de usuarios.

La aplicación de reportes de metadatos (Data Integrator Metadata Reports) provee de reportes sobre el impacto y linaje de datos, tableros operacionales con estadísticas de ejecución de los Jobs y reportes de auto documentación con información crítica sobre los procesos ETL de cada Job de Data Integrator.

El Data Integrator Job Server y el Data Integrator Access Server son manejados por un servicio del sistema operativo. En el sistema operativo Windows, se crea un servicio de nombre "Data Integrator Service", y en el sistema operativo UNIX se crea un demonio de nombre "AL_JobService".

3.4.3 IMPLEMENTACIÓN DE FLUJOS DE CARGA

En el diseñador de Data Integrator, se creó el proyecto DM_SEGUROS_VEHICULOS. Dentro de éste proyecto se crearon los Jobs: JOB_CARGA_DATOS y JOB_CARGA_UNICA. El Job JOB_CARGA_DATOS contiene los flujos de trabajo para transportar los datos de los sistemas fuente al modelo dimensional de datos de forma periódica. El Job JOB_CARGA_UNICA

contiene el flujo de datos para cargar la dimensión fecha que será cargada una única vez.

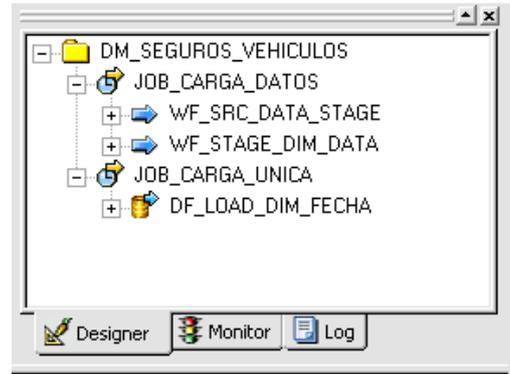


Figura 3.9 Estructura del proyecto DM_SEGUROS_VEHICULOS en Data Integrator Designer

3.4.3.1 Creación de almacenes de datos (Datastores)

Para el desarrollo del proyecto se instaló la base de datos Oracle 10g con la opción Enterprise Edition y se creó la instancia de base de datos de nombre DWH. En el ANEXO A se explican los pasos de instalación del servidor de base de datos y la creación de la instancia de base de datos.

Usando la herramienta de gestión de Oracle, Toad, se crearon los esquemas de Oracle SRC_DATA, STAGE, DIM_DATA y DI_METADATA. Todos los esquemas se crearon usando la siguiente instrucción de Oracle:

```
CREATE USER ESQUEMA IDENTIFIED BY ESQUEMA DEFAULT TABLESPACE DATOS;
GRANT RESOURCE, CONNECT, DBA TO ESQUEMA;
GRANT ALL PRIVILEGES TO ESQUEMA;
```

Como se puede observar el nombre de esquema y la contraseña son las mismas. Además se asignó a todos los esquemas los permisos de DBA y todos los privilegios. Se recomienda por razones de seguridad que las contraseñas de los esquemas creados, en un ambiente de producción, nunca sean las mismas, de igual forma se tiene que decidir que permisos y roles asignar a cada esquema de Oracle. El esquema SRC_DATA se creó usando la siguiente instrucción:

```
CREATE USER SRC_DATA IDENTIFIED BY SRC_DATA DEFAULT TABLESPACE DATOS;
GRANT RESOURCE, CONNECT, DBA TO SRC_DATA;
GRANT ALL PRIVILEGES TO SRC_DATA;
```

De la misma forma se crearon los esquemas de Oracle de nombre STAGE, DIM_DATA y DI_METADATA. El esquema DI_METADATA es el repositorio de la herramienta Data Integrator, en el ANEXO B se explican los pasos de instalación de la herramienta de integración de datos BusinessObjects Data Integrator.

En el diseñador de Data Integrator es necesario crear los almacenes de datos que serán fuente y destino. Como se definió en la arquitectura de flujo de datos se usarán los esquemas o bases de datos SRC_DATA, STAGE y DIM_DATA. Es necesario crear la conexión a éstas bases de datos en el diseñador a través de la opción Datastores de la librería local de objetos. En la Figura 3.10 se ilustra el proceso que se siguió para crear los almacenes de datos para las bases de datos SRC_DATA, STAGE y DIM_DATA.

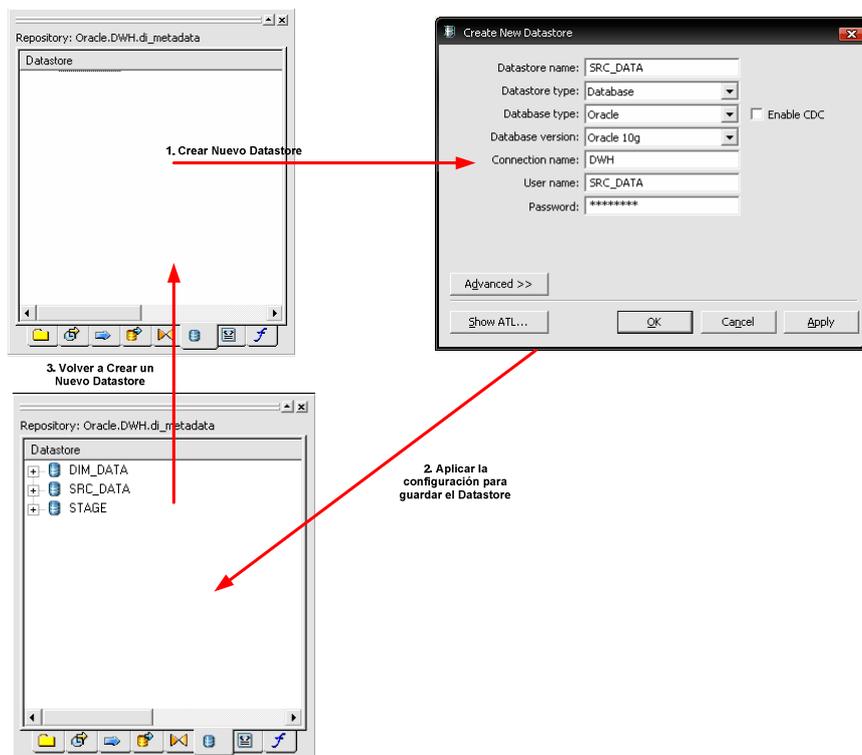


Figura 3.10 Proceso para crear almacenes de datos en Data Integrator Designer

En la opción de datastores se crearon los tres almacenes de datos que se usarán para implementar la arquitectura de flujo de datos del Data Mart de Seguros de Vehículos. Para crear cada Datastore en el diseñador es necesario configurar la base de datos que se usará (Oracle), la versión (Oracle 10g), el nombre de conexión (DWH), el usuario y la contraseña. Se creó los datastores para los esquemas SRC_DATA, STAGE y DIM_DATA siguiendo el mismo proceso.

3.4.3.2 JOB_CARGA_UNICA

El Job de carga único contiene los flujos de trabajo o flujos de datos que se cargarán una única vez. Se creó el flujo de datos DF_LOAD_DIM_FECHA para cargar la dimensión fecha (DIM_FECHA) con fechas que van desde el primero de enero del 2005 hasta el 31 de diciembre del 2015.

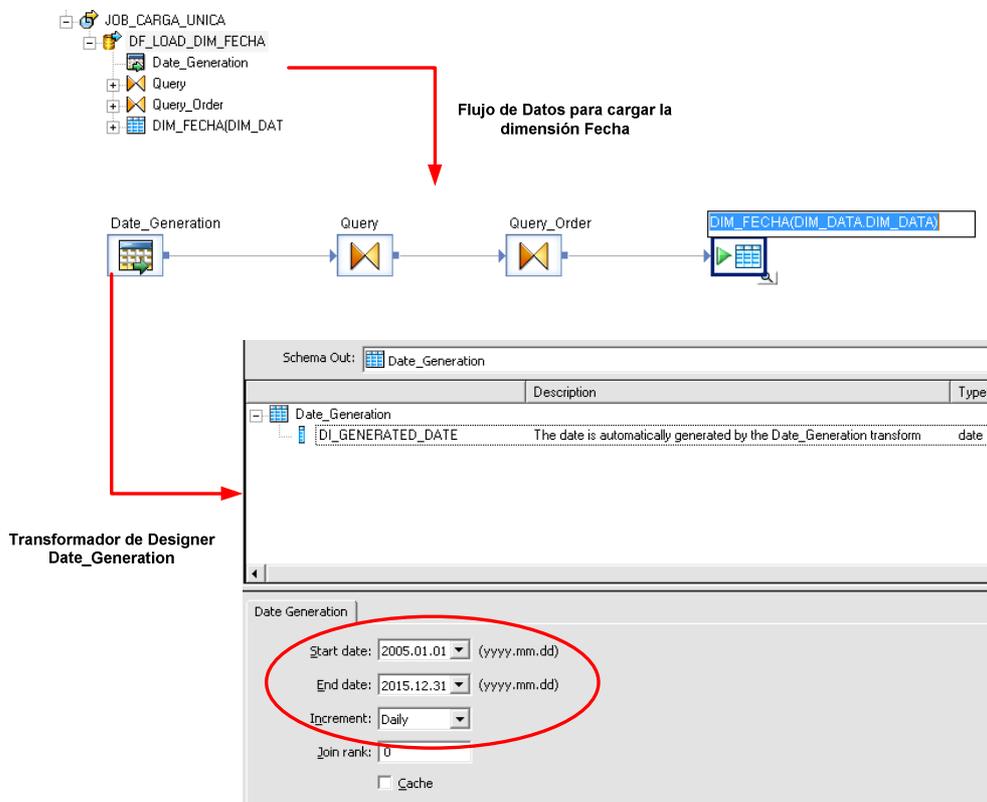


Figura 3.11 Flujo de datos para cargar la dimensión fecha (DF_LOAD_DIM_FECHA)

En la Figura 3.11 se puede observar el flujo de datos creado en el diseñador para cargar la dimensión fecha (DIM_FECHA). El flujo comienza con el transformador de Data Integrator, Date_Generation en el cuál se define la fecha de inicio 2005.01.01, la fecha de fin 2015.12.31 y el incremento. El transformador genera fechas diarias que van desde el primero de enero del 2005 al treinta y uno de diciembre del 2015. En la consulta (Query), se toma la fecha generada por el transformador y se usan funciones del Diseñador para obtener de la fecha generada el año, semestre, trimestre, mes, y día.

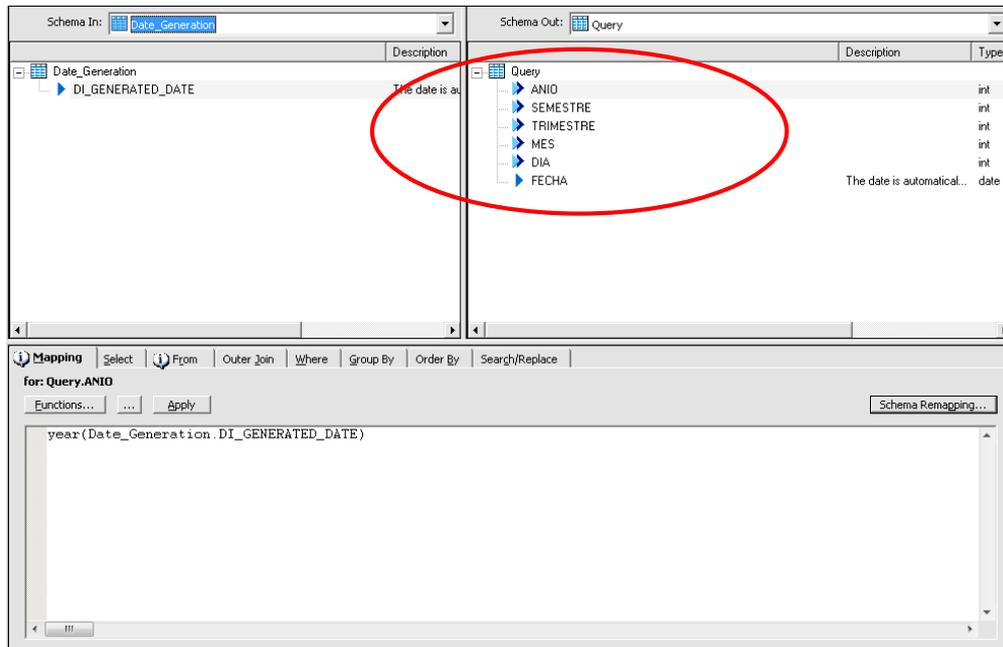


Figura 3.12 Consulta (Query) del flujo de datos DF_LOAD_DIM_FECHA

En la Figura 3.12 se muestra las opciones de configuración del Query del Diseñador. El Query es un transformador que permite crear una sentencia SQL de forma gráfica. Permite definir qué campos de las fuentes tomar en la opción Schema Out, cómo tomar los campos en la opción Mapping, de dónde en la opción From, bajo qué condiciones en la opción Where y si los campos estarán agrupados para funciones de agregación en la opción Group By u ordenados en la opción Order By. En el Query de la dimensión fecha se crearon los siguientes campos de salida:

Campo	Tipo de Dato	Función de Mapeo (Mapping)
ANIO	INT	year(Date_Generation.DI_GENERATED_DATE)
SEMESTRE	INT	DECODE(quarter(Date_Generation.DI_GENERATED_DATE) = 1 OR quarter(Date_Generation.DI_GENERATED_DATE) = 2,1,2)
TRIMESTRE	INT	quarter(Date_Generation.DI_GENERATED_DATE)
MES	INT	month(Date_Generation.DI_GENERATED_DATE)
DIA	INT	day_in_month(Date_Generation.DI_GENERATED_DATE)
FECHA	DATE	Date_Generation.DI_GENERATED_DATE

**Tabla 3.18 Campos generados por el Query del flujo de datos
DF_LOAD_DIM_FECHA**

La consulta Query_Order del flujo de datos DF_LOAD_DIM_FECHA, genera la clave primaria de la dimensión con el nombre SK_FECHA usando la función de mapeo:

```
(Query.ANIO || ifthenelse(Query.MES < 10,'0' || Query.MES,Query.MES) ||  
ifthenelse(Query.DIA < 10,'0' || Query.DIA,Query.DIA))
```

La función de mapeo genera la clave subrogada (SK, Subrogate Key) de la dimensión como un número entero concatenando el año, mes y día obteniéndose una clave de la forma AAAAMMDD. Además, ordena las fechas de forma ascendente por año, semestre, trimestre, mes y día para cargar la dimensión DIM_FECHA del Data Mart que físicamente se encuentra en el esquema DIM_DATA. En el ANEXO C se describe el proceso de creación de la base de datos dimensional (DIM_DATA) que contiene las dimensiones y tablas de hechos del Data Mart de seguros de vehículos.

En el diseñador de Data Integrator se puede ejecutar el Job (JOB_CARGA_UNICA) para ejecutar el flujo de datos (DF_LOAD_DIM_FECHA) y cargar la dimensión DIM_FECHA. En la Figura 3.13 se muestra la dimensión fecha cargada con datos después de ejecutar el Job de carga única.

SK_FECHA	ANIO	SEMESTRE	TRIMESTRE	MES	DIA	FECHA
20050101	2005	1	1	1	1	2005.01.01 00:00:00.000000000
20050102	2005	1	1	1	2	2005.01.02 00:00:00.000000000
20050103	2005	1	1	1	3	2005.01.03 00:00:00.000000000
20050104	2005	1	1	1	4	2005.01.04 00:00:00.000000000
20050105	2005	1	1	1	5	2005.01.05 00:00:00.000000000
20050106	2005	1	1	1	6	2005.01.06 00:00:00.000000000
20050107	2005	1	1	1	7	2005.01.07 00:00:00.000000000
20050108	2005	1	1	1	8	2005.01.08 00:00:00.000000000
20050109	2005	1	1	1	9	2005.01.09 00:00:00.000000000
20050110	2005	1	1	1	10	2005.01.10 00:00:00.000000000
20050111	2005	1	1	1	11	2005.01.11 00:00:00.000000000
20050112	2005	1	1	1	12	2005.01.12 00:00:00.000000000
20050113	2005	1	1	1	13	2005.01.13 00:00:00.000000000

Figura 3.13 Dimensión Fecha (DIM_FECHA) cargada de datos

3.4.3.3 JOB_CARGA_DATOS

El Job JOB_CARGA_DATOS contiene los flujos de trabajo (Work Flows) para transportar los datos del esquema SRC_DATA al esquema STAGE y del esquema STAGE al esquema DIM_DATA. Los flujos de trabajo, contienen los flujos de datos (Data Flows) que realizan los procesos ETL para cargar cada dimensión del modelo dimensional y las tablas de hechos de producción y siniestros. En la Figura 3.14 se muestra la estructura del Job JOB_CARGA_DATOS:



Figura 3.14 Estructura de los flujos de trabajo del JOB_CARGA_DATOS

3.4.3.3.1 Flujo de Trabajo WF_SRC_DATA_STAGE

El flujo de trabajo WF_SRC_DATA_STAGE contiene flujos de datos que extraen los datos de las tablas fuente del esquema SRC_DATA, transforman los datos y los depositan limpios en tablas temporales del esquema STAGE.

ETL_COBERTURA

El flujo de trabajo ETL_COBERTURA contiene los flujos de datos DF_LOAD_COBERTURA_001 y DF_LOAD_COBERTURA_002 para transportar los datos de coberturas al esquema STAGE.

En el flujo de datos DF_LOAD_COBERTURA_001 se crea una nueva clave que se convertirá en la clave subrogada de la dimensión cobertura. En este flujo de datos se transportan los registros de la tabla fuente COBERTURA hacia la tabla temporal MAPEO_COBERTURA generándose una nueva clave entera por cada cobertura.

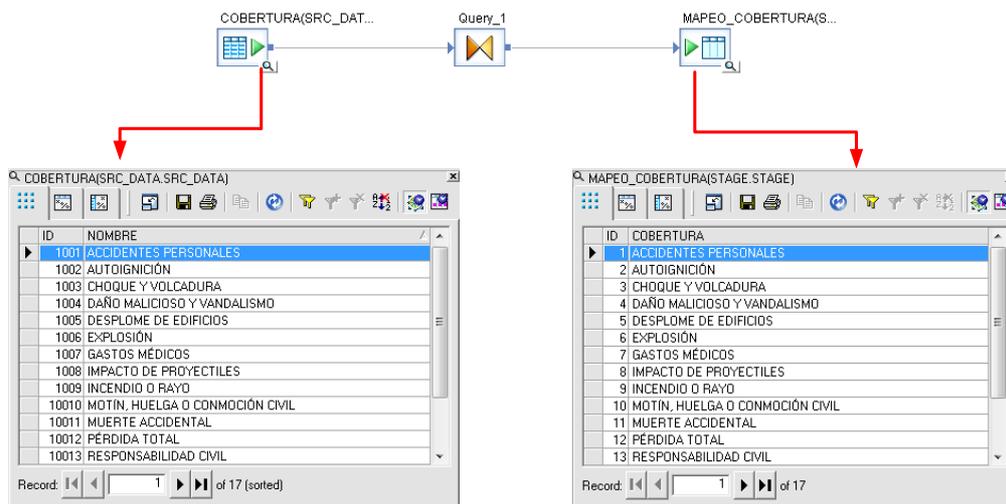


Figura 3.15 Flujo de datos DF_LOAD_COBERTURA_001

Como se puede observar en la Figura 3.15 el flujo de datos DF_LOAD_COBERTURA_001 en el Query_1 crea un ID único usando la función `gen_row_num()` para cada nombre de cobertura único. En la tabla temporal MAPEO_COBERTURA del esquema STAGE se carga un registro de cobertura con un nuevo ID como se puede observar en la Figura 3.16.

ID	NOMBRE
1001	ACCIDENTES PERSONALES
1002	AUTOIGNICIÓN
1003	CHOQUE Y VOLCADURA
1004	DAÑO MALICIOSO Y VANDALISMO
1005	DESPLOME DE EDIFICIOS
1006	EXPLOSIÓN
1007	GASTOS MÉDICOS
1008	IMPACTO DE PROYECTILES
1009	INCENDIO O RAYO
10010	MOTÍN, HUELGA O CONMOCIÓN CIVIL
10011	MUERTE ACCIDENTAL
10012	PÉRDIDA TOTAL
10013	RESPONSABILIDAD CIVIL

ID	COBERTURA
1	ACCIDENTES PERSONALES
2	AUTOIGNICIÓN
3	CHOQUE Y VOLCADURA
4	DAÑO MALICIOSO Y VANDALISMO
5	DESPLOME DE EDIFICIOS
6	EXPLOSIÓN
7	GASTOS MÉDICOS
8	IMPACTO DE PROYECTILES
9	INCENDIO O RAYO
10	MOTÍN, HUELGA O CONMOCIÓN CIVIL
11	MUERTE ACCIDENTAL
12	PÉRDIDA TOTAL
13	RESPONSABILIDAD CIVIL

Figura 3.16 Carga de la tabla temporal MAPEO_COBERTURA

El flujo de datos DF_LOAD_COBERTURA_002 realiza una juntura entre la tabla temporal MAPEO_COBERTURA y la tabla COBERTURA para crear una tabla de mapeo auxiliar que contenga el identificador original de la tabla COBERTURA y el nuevo identificador generado, esto se puede observar en la Figura 3.17.

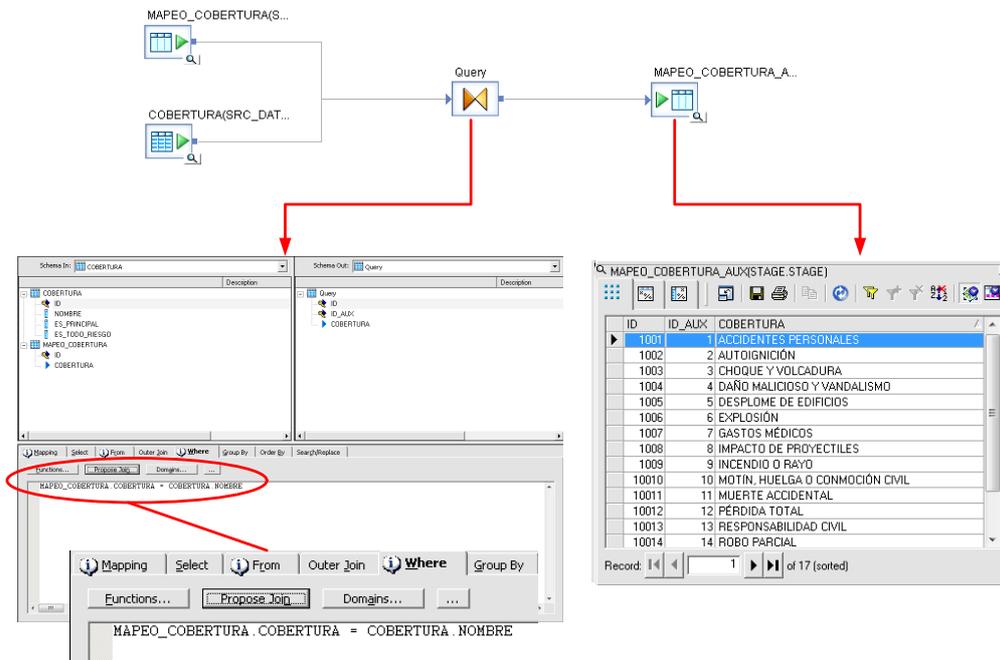


Figura 3.17 Flujo de datos DF_LOAD_COBERTURA_002

ETL_AGENCIA

El flujo de trabajo ETL_AGENCIA contiene el flujo de datos DF_LOAD_AGENCIA_ENCARGADO el cual lee de un archivo de texto el nombre del encargado por cada agencia. El nombre del encargado no se encontraba

como campo de la tabla fuente AGENCIA, por lo que se requería cargarlo desde el archivo de texto AGENCIA_ENCARGADO.txt a la tabla temporal AGENCIA_ENCARGADO, como se muestra en la Figura 3.18.

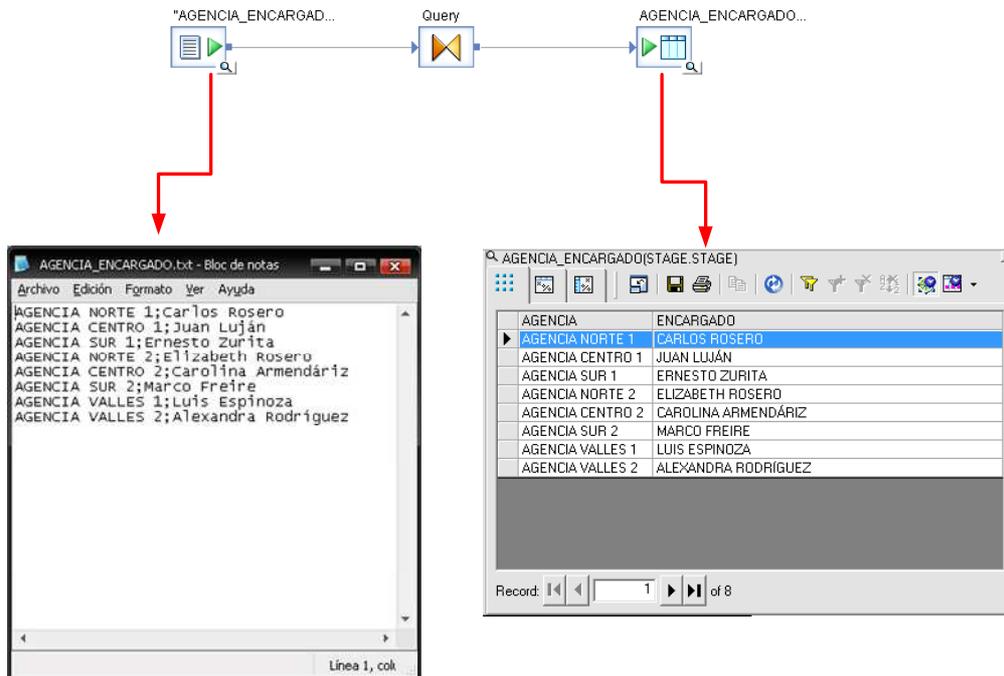


Figura 3.18 Flujo de datos DF_LOAD_AGENCIA_ENCARGADO

ETL_ASEGURADO

El flujo de trabajo ETL_ASEGURADO contiene los flujos de datos DF_LOAD_ASEGURADO_001 y DF_LOAD_ASEGURADO_002 para transportar los datos de asegurados al esquema STAGE.

De igual forma que para la tabla COBERTURA, se encontró que para la tabla ASEGURADO existía redundancia de nombres de asegurados por cada ENDOSOITEMID. Se implementó un proceso de carga similar al realizado en flujo de trabajo ETL_COBERTURA.

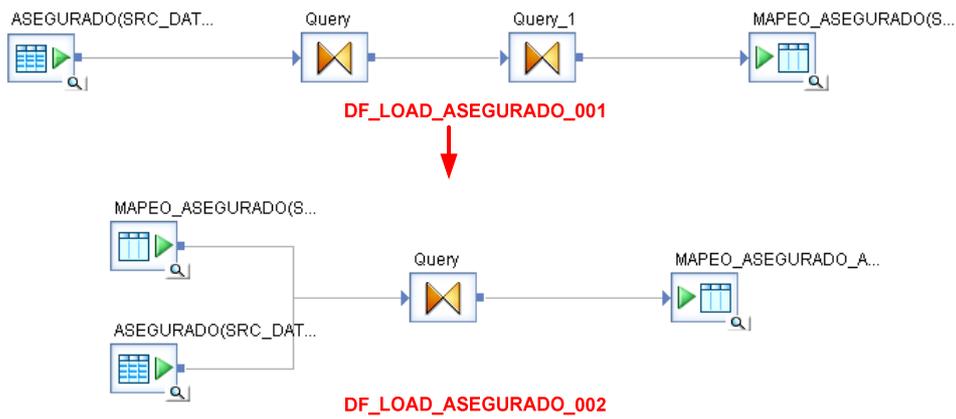


Figura 3.19 Flujos de datos del flujo de trabajo ETL_ASEGURADO

Al final de la carga de los flujos de datos DF_LOAD_ASEGURADO_001 y DF_LOAD_ASEGURADO_002 se ingresa en las tablas temporales MAPEO_ASEGURADO_AUX y MAPEO_ASEGURADO un registro de asegurado desconocido para controlar más adelante posibles problemas de falta de integridad referencial. Se usó la función `sql()` de la siguiente forma:

```
sql('STAGE', 'INSERT INTO MAPEO_ASEGURADO_AUX VALUES(\ '0\ ', 0, \ 'DESCONOCIDO\ ');
sql('STAGE', 'INSERT INTO MAPEO_ASEGURADO VALUES (0, \ 'DESCONOCIDO\ ');
```

ETL_BENEFICIARIO

El flujo de trabajo ETL_BENEFICIARIO contiene los flujos de datos DF_LOAD_BENEFICIARIO_001 y DF_LOAD_BENEFICIARIO_002 para transportar los datos de beneficiarios al esquema STAGE. En éste flujo de trabajo se tuvo el mismo problema que el flujo de trabajo ETL_ASEGURADO, y se implementó de forma similar como se muestra en la Figura 3.20.

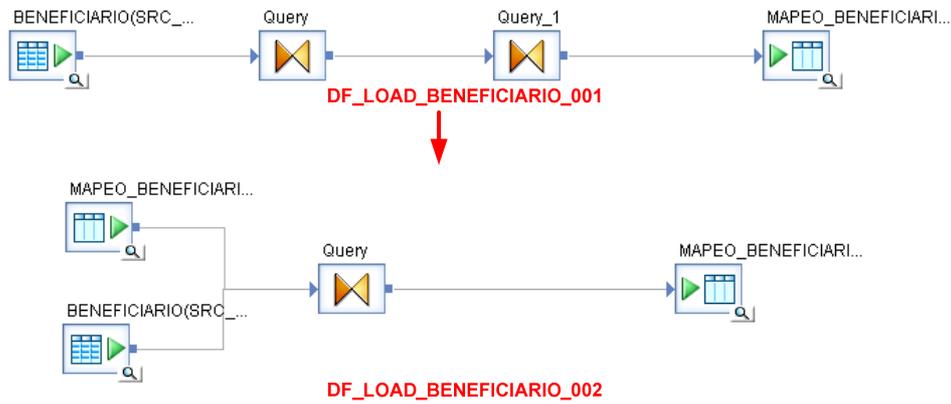


Figura 3.20 Flujos de datos del flujo de trabajo ETL_BENEFICIARIO

Después de la carga de los flujos DF_LOAD_BENEFICIARIO_001 y 002 se insertó un registro de desconocido para controlar la pérdida de integridad en las cargas posteriores sobre las tablas temporales MAPEO_BENEFICIARIO y MAPEO_BENEFICIARIO_AUX. Se utilizó la función `sql()` a través de un script del Diseñador de Data Integrator de la siguiente forma:

```
sql('STAGE','INSERT INTO MAPEO_BENEFICIARIO_AUX VALUES (\ '0\ ', 0,
\ 'DESCONOCIDO\ ')');
sql('STAGE','INSERT INTO MAPEO_BENEFICIARIO VALUES (0,\ 'DESCONOCIDO\ ')');
```

ETL_EXTRA_ASEGURADO

El flujo de trabajo ETL_EXTRA_ASEGURADO contiene los flujos de datos DF_CLEAN_EXTRA_ASEGURADOS_001 y DF_CLEAN_EXTRA_ASEGURADOS_002 para limpiar la descripción del extra, utilizando un nombre general para cada descripción. Por ejemplo, si un extra tenía como descripción EQUIPO KENWOOD y otro EQUIPO SONY a los dos se los nombró con el nombre general EQUIPO DE MÚSICA. Este mapeo de descripción versus el nombre general se lo creó a través de un archivo plano. Los flujos de datos DF_LOAD_EXTRA_ASEGURADO_003 y DF_LOAD_EXTRA_ASEGURADO_004 se encargan de transportar únicamente los nombres generales del extra asegurado, generando nuevas claves y creando una tabla de mapeo que preserve la clave original del extra y la nueva clave generada para el nombre general del extra asegurado.

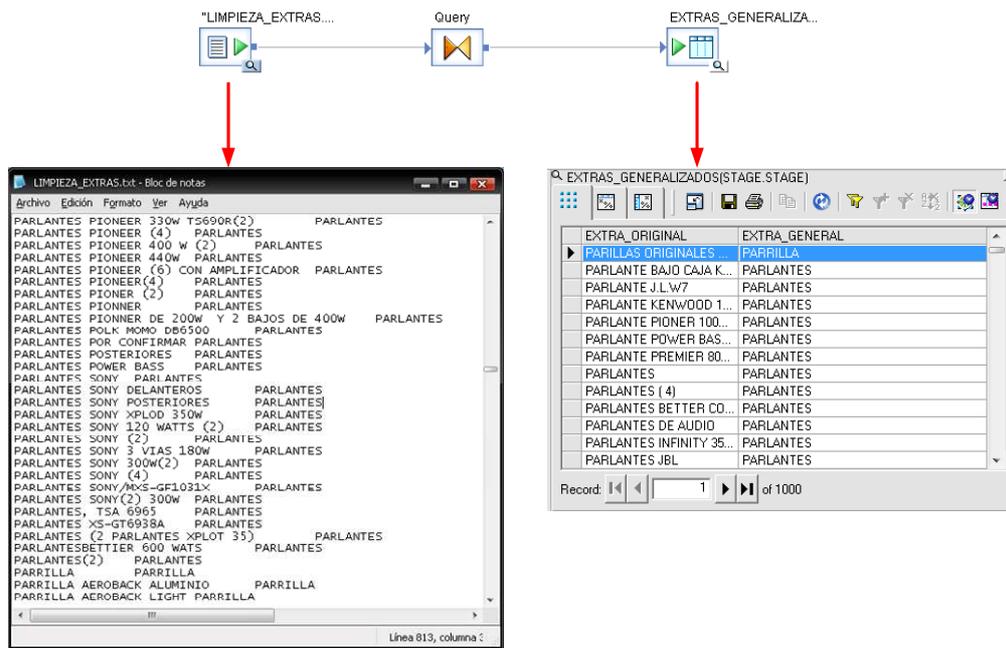


Figura 3.21 Flujo de datos DF_CLEAN_EXTRA_ASEGURADOS_001

En el flujo de datos DF_CLEAN_EXTRA_ASEGURADOS_001 se carga el archivo plano separado por tabulaciones LIMPIEZA_EXTRAS.txt a la tabla temporal EXTRAS_GENERALIZADOS. Como se puede observar cada descripción del extra contiene un nombre de extra generalizado.

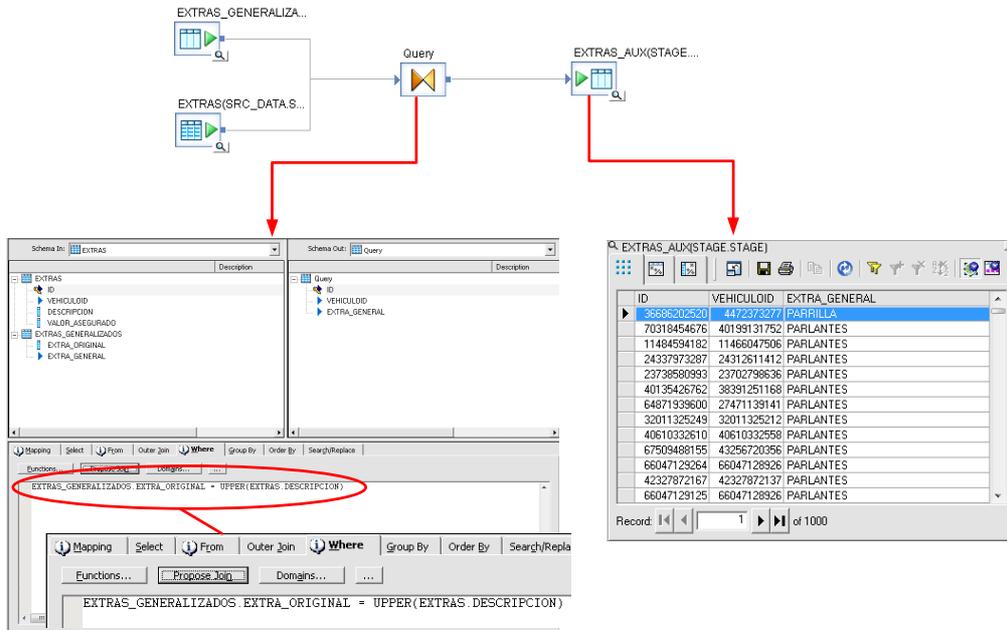


Figura 3.22 Flujo de datos DF_CLEAN_EXTRA_ASEGURADOS_001

El flujo de datos DF_CLEAN_EXTRA_ASEGURADOS_001 cruza la tabla EXTRAS_GENERALIZADOS y la tabla fuente EXTRAS para carga la tabla temporal EXTRAS_AUX que contiene el ID del extra, el ID del vehículo y el nombre del extra generalizado.

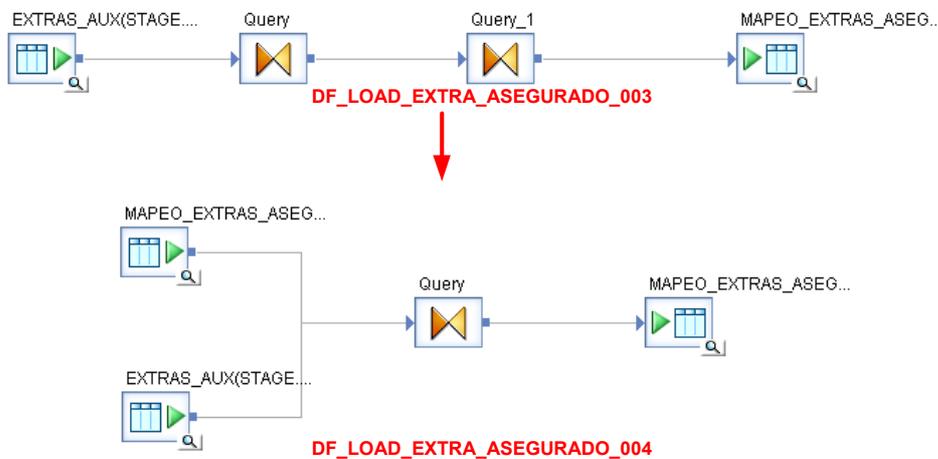


Figura 3.23 Flujo de datos DF_LOAD_EXTRA_ASEGURADO_003 y 004

El flujo de datos DF_LOAD_EXTRA_ASEGURADO_003 extrae los datos distintos del nombre del extra generalizado de la tabla EXTRAS_AUX y crea un nuevo

identificador por cada nombre de extra generalizado único cargando la tabla temporal MAPEO_EXTRAS_ASEGURADOS. El flujo de datos DF_LOAD_EXTRA_ASEGURADO_004 cruza la tabla temporal generada con la tabla EXTRAS_AUX para obtener una tabla de mapeo que contenga el identificador generado por cada extra generalizado y el identificador original de cada extra. Como se puede observar, se aplica el mismo método de carga implementado en los flujos de trabajo ETL_COBERTURA, ETL_ASEGURADO y ETL_BENEFICIARIO.

ETL_VEHICULO

El flujo de trabajo ETL_VEHICULO contiene los flujos de trabajo DF_CLEAN_VEHICULO_001, DF_CLEAN_VEHICULO_002 y DF_CLEAN_VEHICULO_003 para realizar tareas de limpieza de los datos necesarios de vehículos para cargar la dimensión DIM_VEHICULO.

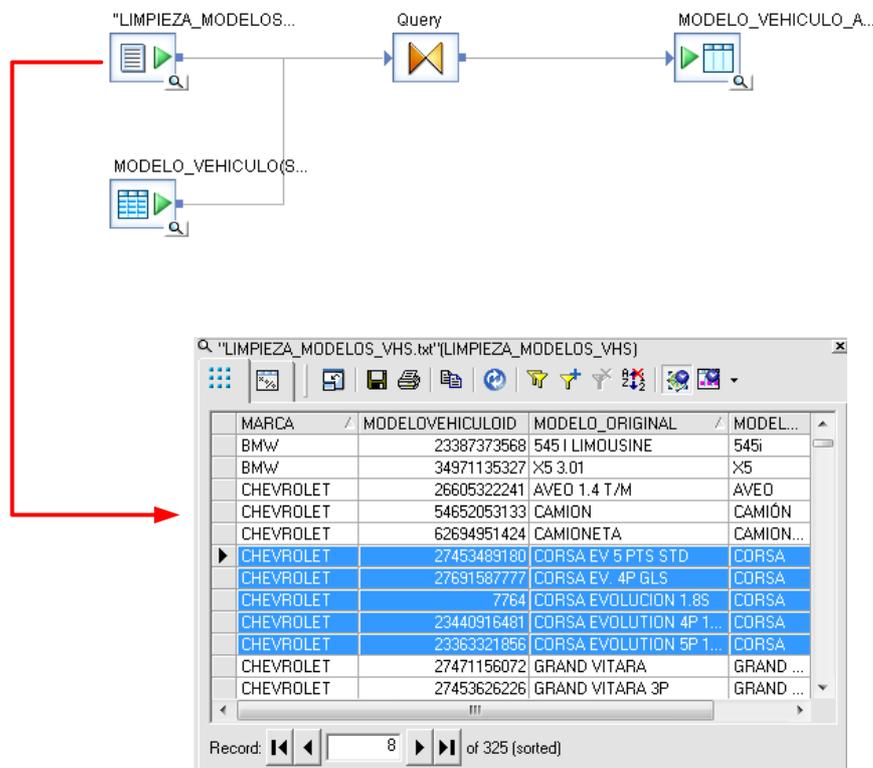


Figura 3.24 Flujo de datos DF_CLEAN_VEHICULO_001

Entre los procesos de limpieza se realizó una generalización de los modelos de vehículos por ejemplo para un modelo de vehículo CORSA EVOLUCION 4P y un

modelo CORSA EVOLUTION 5P se creó un nombre de modelo generalizado CORSA. Así se generalizó para todos los modelos creando el archivo de mapeo de nombre LIMPIEZA_MODELOS_VHS.txt con el nombre del modelo original y el nombre de modelo generalizado. El archivo plano se lo cargó en la tabla MODELO_VEHICULO_AUX del esquema STAGE como se puede observar en la Figura 3.24.

En el flujo de datos DF_CLEAN_VEHICULO_002 se cargaron los datos de marcas de vehículos, realizando el proceso de limpieza de la marca INTERNACIONAL colocándola como INTERNATIONAL, como se muestra en la Figura 3.25.

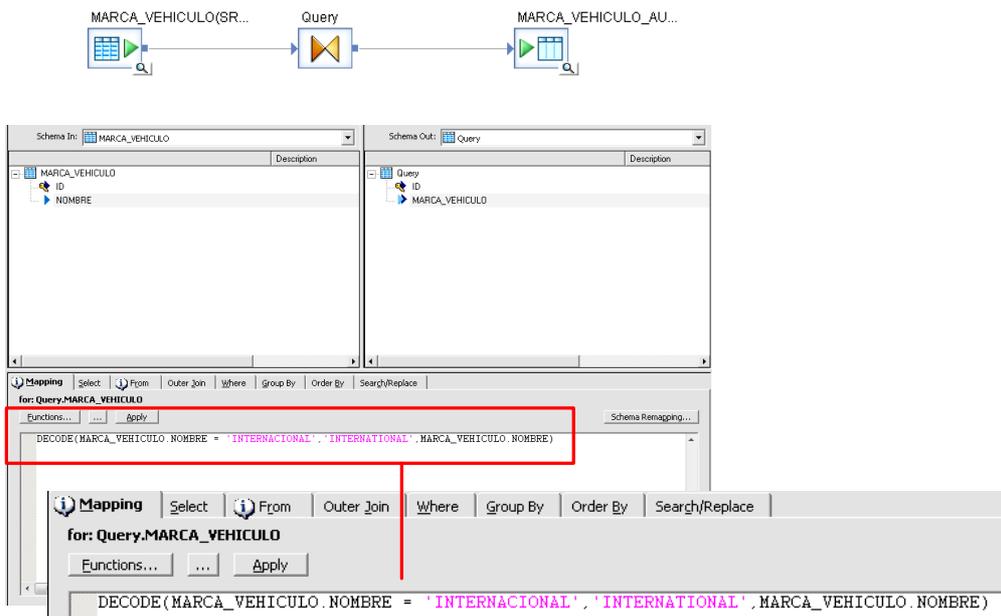


Figura 3.25 Flujo de datos DF_CLEAN_VEHICULO_002

Como se puede observar se utilizó la función DECODE para limpiar la marca mencionada. En el flujo de datos DF_CLEAN_VEHICULO_003 se cruzó las tablas fuente ITEM, VEHICULO y ZONA_RIESGO, como se puede observar en la

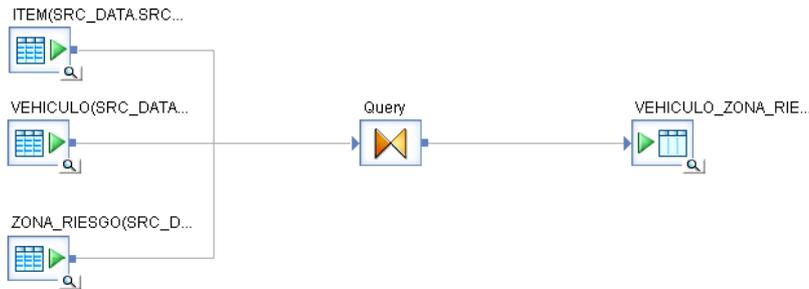


Figura 3.26 Flujo de datos DF_LOAD_VEHICULO_003

En la opción de mapeo del Query creado en el flujo de datos DF_LOAD_VEHICULO_003 se realizaron las siguientes transformaciones de datos descritas en la Tabla 3.19.

Campo	Tipo de Dato	Función de Mapeo (Mapping)
PLACAS	VARCHAR (10)	ifthenelse(length(VEHICULO.PLACAS) IN (6,7),UPPER(VEHICULO.PLACAS) , 'DDD0000')
MOTOR	VARCHAR (40)	ifthenelse(length(VEHICULO.MOTOR) < 5, 'DDD0000' ,UPPER(VEHICULO.MOTOR))
CHASIS	VARCHAR (40)	ifthenelse(length(VEHICULO.CHASIS) < 5, 'DDD0000' ,UPPER(VEHICULO.CHASIS))
NUMERO_O CUPANTES	INT	ifthenelse(VEHICULO.NUMERO_OCUPANTES = 0 OR VEHICULO.NUMERO_OCUPANTES IS NULL,1,VEHICULO.NUMERO_OCUPANTES)
ANIO_FABR ICACION	INT	ifthenelse(VEHICULO.ANIO_FABRICACION < 1970,1970,ifthenelse(VEHICULO.ANIO_FABRICACION > (year(sysdate()) + 1),(year(sysdate()) + 1),VEHICULO.ANIO_FABRICACION))

Tabla 3.19 Campos generados en el Query del flujo de datos DF_LOAD_VEHICULO_003

Los datos desconocidos de números de placas, motor y chasis se los colocó como DDD0000. Se definió que el número mínimo de ocupantes sea 1 y que los años de fabricación nulos sean mayores que 1970 y un año más que el año en curso. Después de haber realizado las tareas de limpieza de datos de vehículos se cargó la tabla temporal VEHICULO_AUX en el esquema STAGE a través del flujo de datos DF_LOAD_VEHICULO_004 como se muestra en la Figura 3.27.

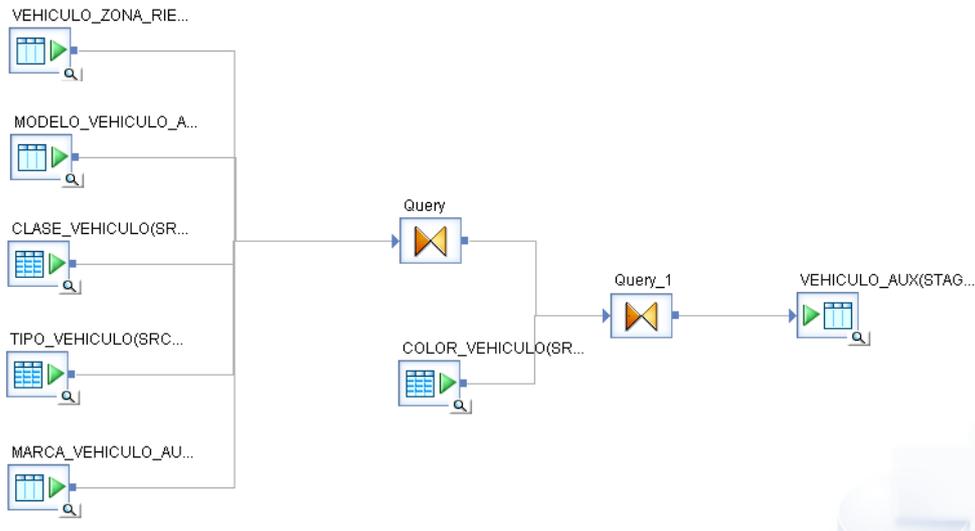


Figura 3.27 Flujo de datos DF_LOAD_VEHICULO_004

ETL_CLIENTE

La carga de la tabla temporal con los datos de los clientes se la realizó a través de un único flujo de datos de nombre DF_LOAD_CLIENTE_001. Las actividades de limpieza que se realizaron fueron la eliminación de espacios en blanco en los nombres de clientes y el control del tipo de cliente analizando la bandera que indicaba sí el cliente era empresa o no.

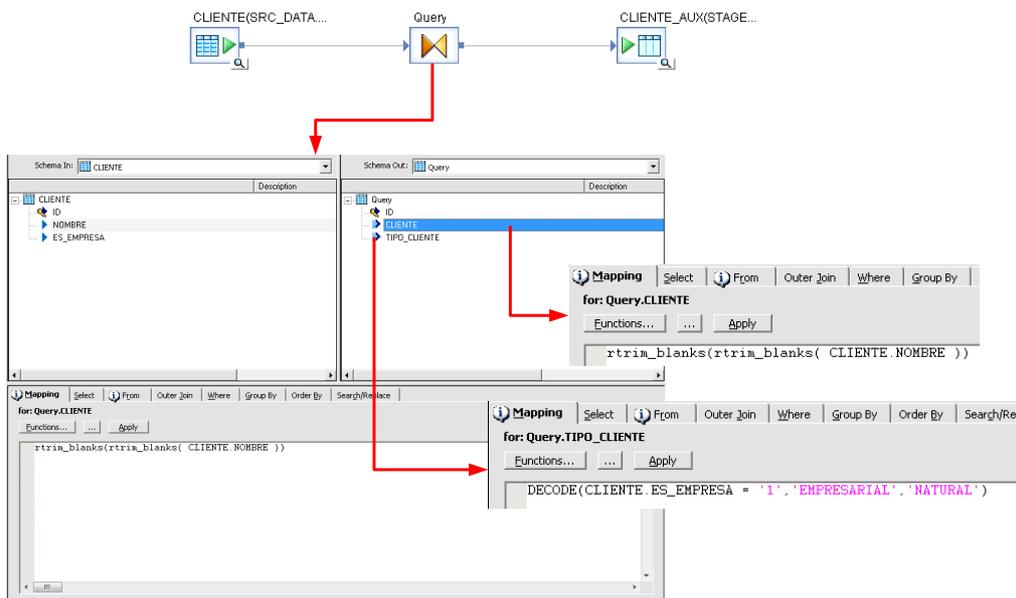


Figura 3.28 Flujo de datos DF_LOAD_CLIENTE_001

ETL_ENDOSO

El flujo de trabajo ETL_ENDOSO contiene los flujos de datos DF_LOAD_ENDOSO_001 y DF_LOAD_ENDOSO_002 para cargar los datos de endosos o anexos de pólizas. El flujo de datos DF_LOAD_ENDOSO_001 cruza las tablas fuente con los datos de endosos necesarios, tomando únicamente aquellos endosos que tiene el estado ES_ACTUAL como 1.

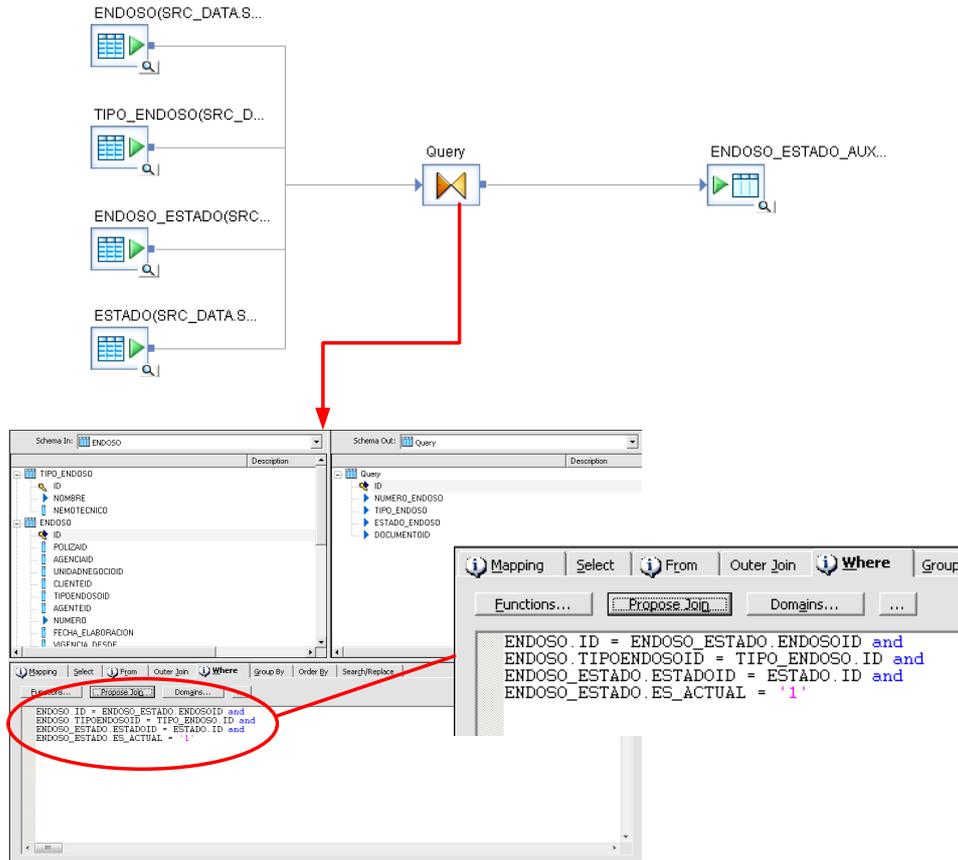


Figura 3.29 Flujo de datos DF_LOAD_ENDOSO_001

El flujo de datos DF_LOAD_ENDOSO_001 almacena los datos de endosos en la tabla temporal ENDOSO_ESTADO_AUX del esquema STAGE. El flujo de datos DF_LOAD_ENDOSO_002 cruza los endosos con los documentos a los cuales está asociado cada estado del endoso. Los documentos pueden ser facturas o notas de crédito.

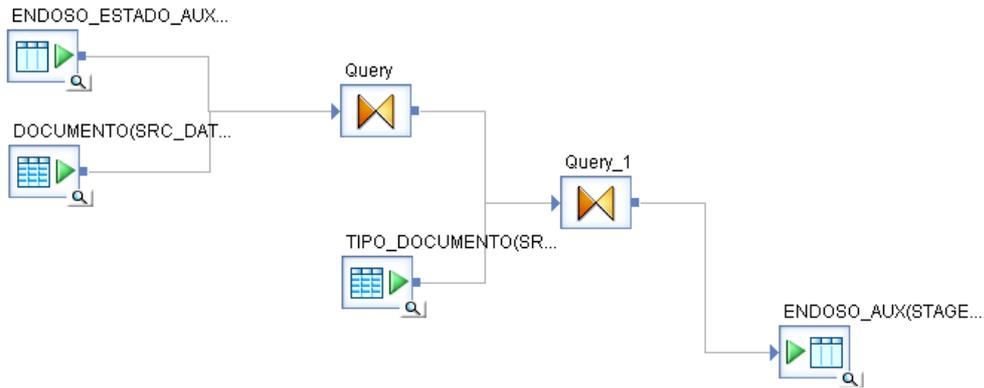


Figura 3.30 Flujo de datos DF_LOAD_ENDOSO_002

En la carga de datos se encontró que faltaba integridad referencial entre la tabla fuente documento y la tabla temporal. También se encontró que faltaba integridad referencial entre la tabla fuente DOCUMENTO y TIPO_DOCUMENTO. Fue necesario tanto en el Query como en el Query_1 realizar un OUTER JOIN para solucionar el problema de integridad referencial.

ETL_PRODUCCION_POLIZAS

El flujo de trabajo ETL_PRODUCCION_POLIZAS se encarga de cargar la tabla temporal de hechos de producción de pólizas. El flujo de trabajo contiene el flujo de trabajo ETL_PRODUCCION_PRIMA_SUSCRITA para cargar las primas suscritas y el flujo de trabajo ETL_PRODUCCION_PRIMA_COBRADA para cargar las primas cobradas.

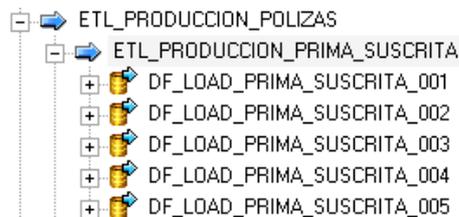


Figura 3.31 Flujo de trabajo ETL_PRODUCCION_PRIMA_SUSCRITA

El flujo de trabajo ETL_PRODUCCION_PRIMA_SUSCRITA se encarga de cargar los endosos suscritos y obtener el valor de prima suscrita. El flujo de datos DF_LOAD_PRIMA_SUSCRITA_001 carga los endosos con estado 15 (FACTURADO) y toma el mínimo orden de la tabla fuente ENDOSO_ESTADO, como se puede observar en la Figura 3.32.

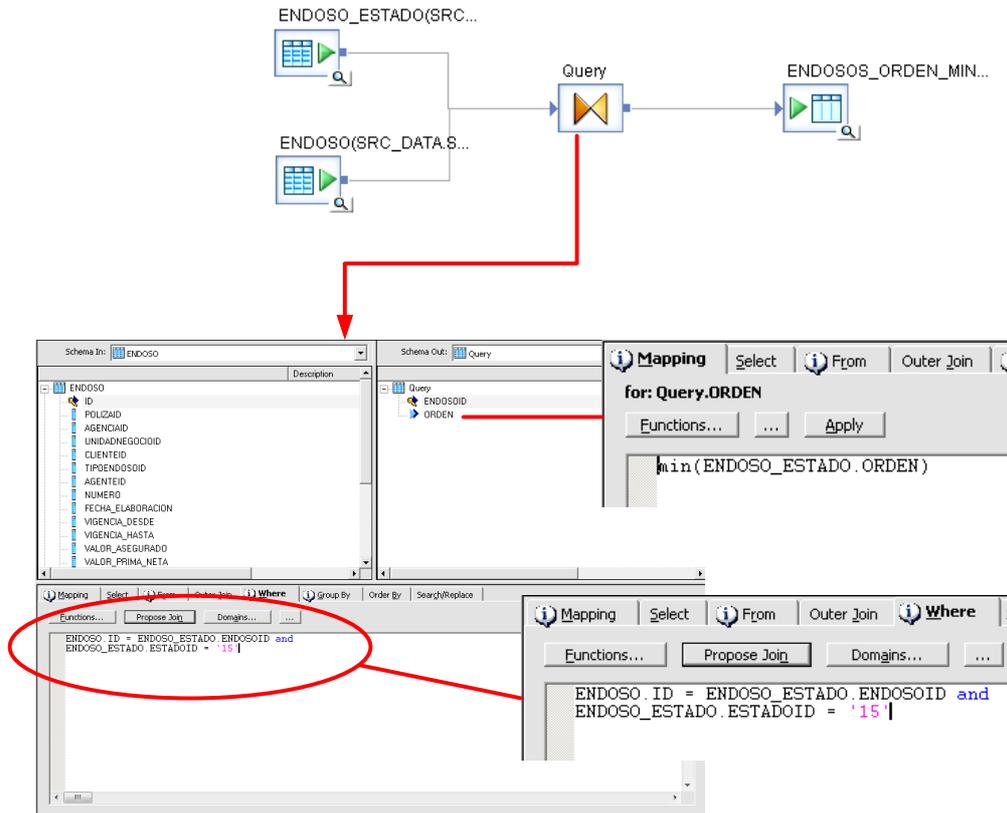


Figura 3.32 Flujo de datos DF_LOAD_PRIMA_SUSCRITA_001

El flujo de datos DF_LOAD_PRIMA_SUSCRITA_002 carga los endosos suscritos positivos que han sido facturados, estado 15 y no anulados, estado 14. Además toma el último estado del endoso con la condición que el endoso sea actual. Este flujo de datos es dependiente del flujo de datos DF_LOAD_PRIMA_SUSCRITA_001 como se muestra en la Figura 3.33.

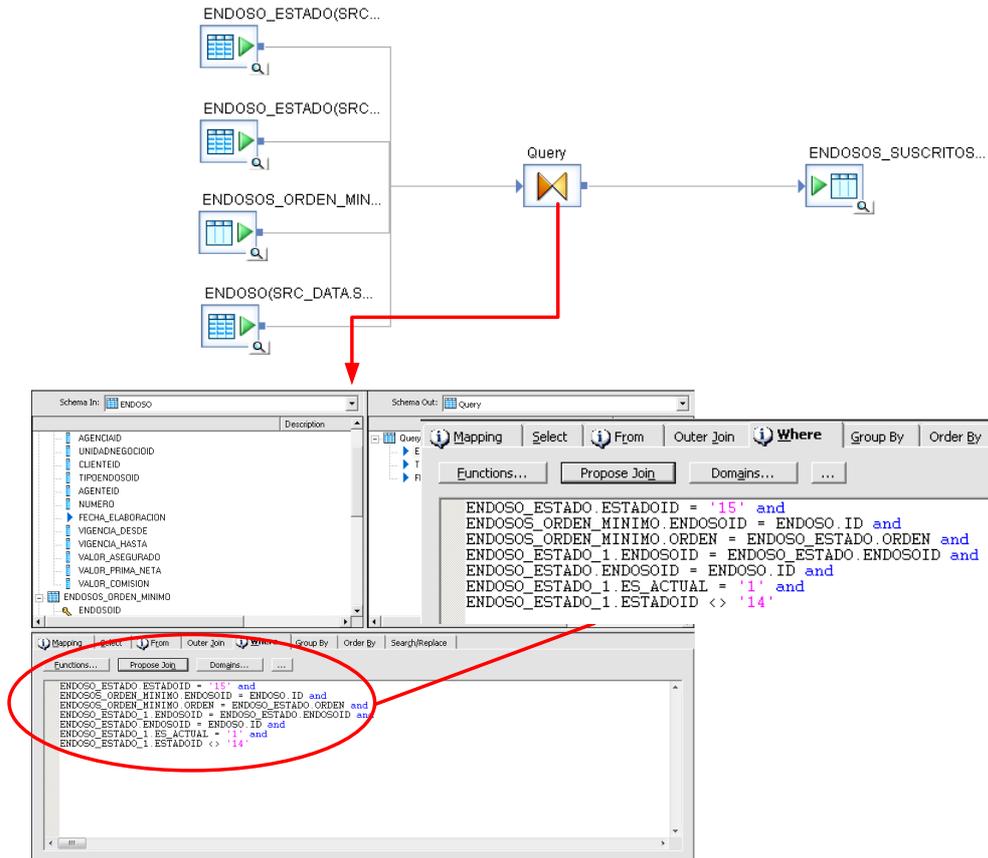


Figura 3.33 Flujo de datos DF_LOAD_PRIMA_SUSCRITA_002

El flujo de datos DF_LOAD_PRIMA_SUSCRITA_003 carga los endosos suscritos negativos. Carga aquellos endosos que fueron pagados pero su valor de primas es negativo por cancelaciones de pólizas. Además, este flujo de datos carga el último estado del endoso con la condición verdadera (1) del campo ES_ACTUAL.

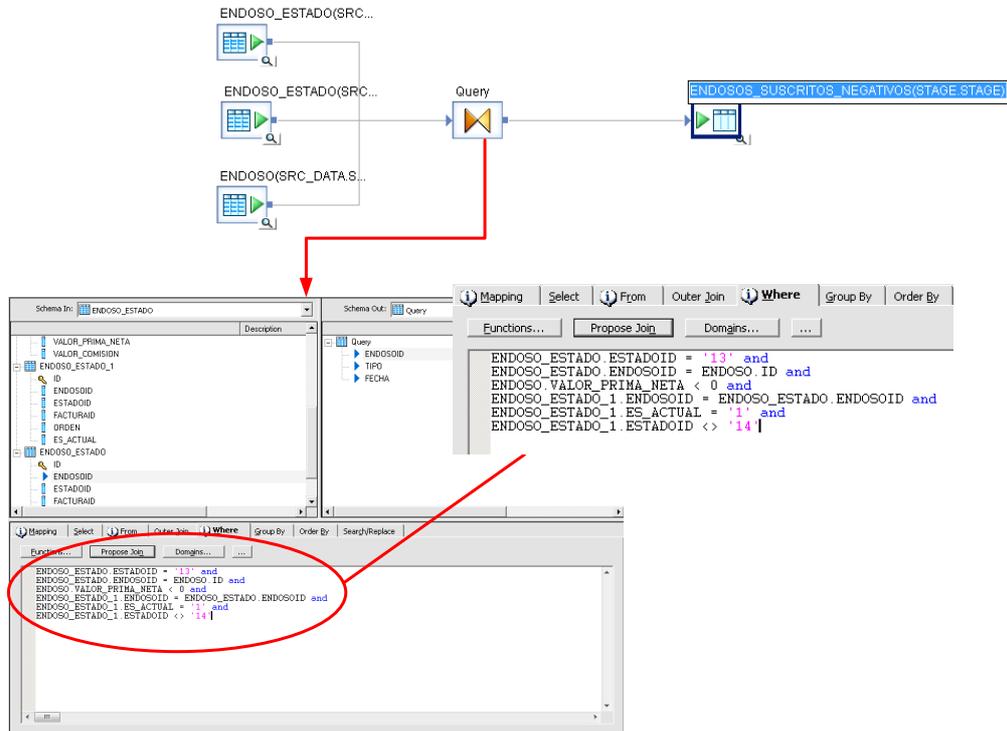


Figura 3.34 Flujo de datos DF_LOAD_PRIMA_SUSCRITA_003

El flujo de datos DF_LOAD_PRIMA_SUSCRITA_004 hace una unión de las tablas temporales `ENDOSOS_SUSCRITOS_POSITIVOS` y `ENDOSOS_SUSCRITOS_NEGATIVOS` en una única tabla temporal de nombre `ENDOSOS_SUSCRITOS` dentro del esquema `STAGE`. En el flujo de datos se usa el transformado `Merge` para realizar unión entre tablas.

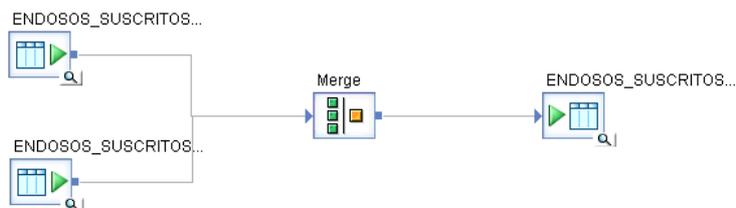


Figura 3.35 Flujo de datos DF_LOAD_PRIMA_SUSCRITA_004

El flujo de datos DF_LOAD_PRIMA_SUSCRITA_005 cruza la tabla temporal de `ENDOSOS_SUSCRITOS` con las tablas fuente de dónde se obtendrá las claves naturales de las futuras dimensiones. Además, se extraen los campos que se

cargarán en la tabla de hechos FACT_PRODUCCION como se puede observar en la Figura 3.36.

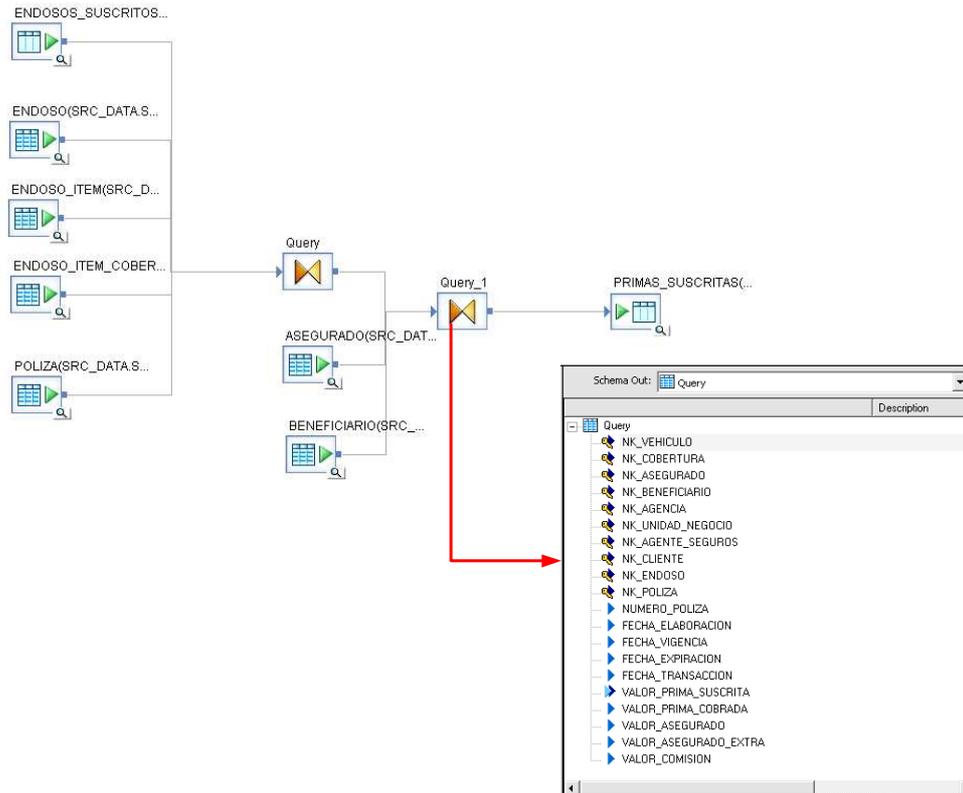


Figura 3.36 Flujo de datos DF_LOAD_PRIMA_SUSCRITA_005

El flujo de trabajo ETL_PRODUCCION_PRIMA_COBRADA contiene los flujos de datos para cargar las primas cobradas. Las primas cobradas son aquellas que han sido pagadas y han generado un asiento contable. Los flujos de datos cargan las primas cobradas positivas, primas cobradas negativas y las primas cobradas por concepto de extras asegurados. En la Figura 3.37 se puede observar la estructura del flujo de trabajo ETL_PRODUCCION_PRIMA_COBRADA y los flujos de datos definidos.

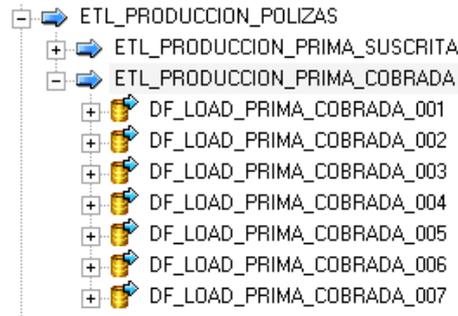


Figura 3.37 Flujo de trabajo ETL_PRODUCCION_PRIMA_COBRADA

El flujo de datos DF_LOAD_PRIMA_COBRADA_001 cruza las tablas fuentes que contienen las claves naturales de las dimensiones y toma los endosos que se encuentran en estado 15 (FACTURADO). Además, el flujo de datos extrae el último estado del endoso y las coberturas que afectan prima y valor asegurado como se muestra en la Figura 3.38. Los datos son cargados en la tabla temporal PRIMA_COBRADA_POS_AUX del esquema STAGE.

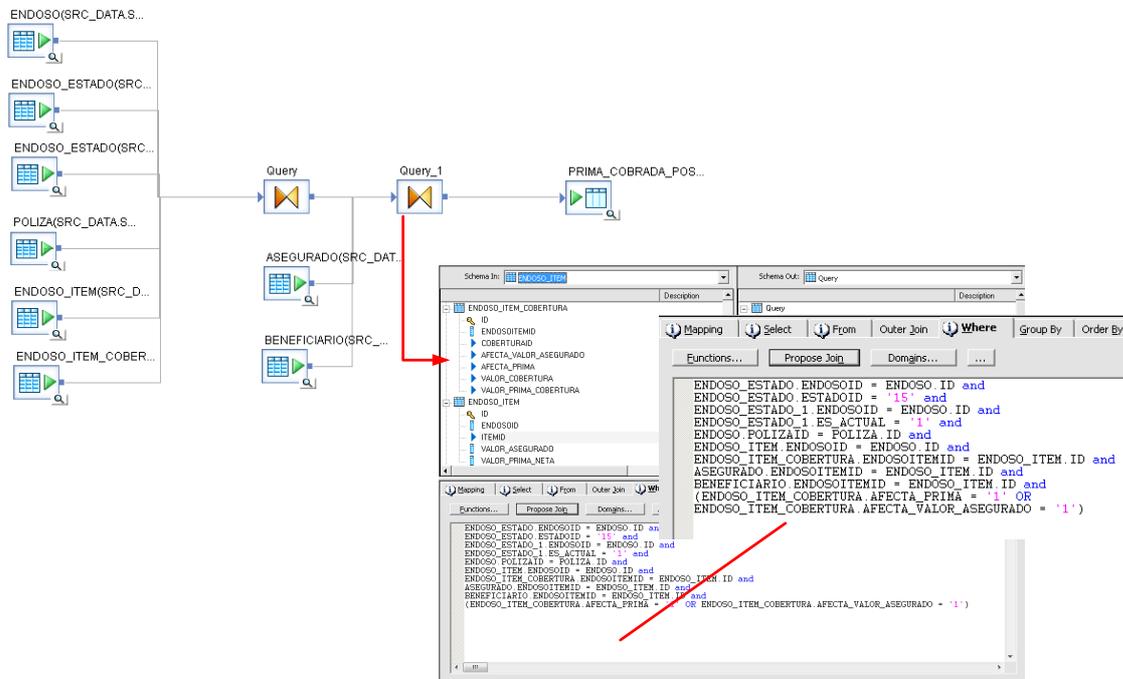


Figura 3.38 Flujo de datos DF_LOAD_PRIMA_COBRADA_001

El flujo de datos DF_LOAD_PRIMA_COBRADA_002 cruza la tabla temporal PRIMA_COBRADA_POS_AUX con la tabla fuente DOCUMENTO extrayendo los documentos que no estén anulados y sean facturas de clientes. En el Query_1 se

cruzan los documentos con los asientos para extraer sólo aquellos documentos que han sido contabilizados tomando el estado 301 de la tabla fuente ASIENTO. Los datos son cargados en la tabla temporal PRIMA_COBRADA_POSITIVA del esquema STAGE.

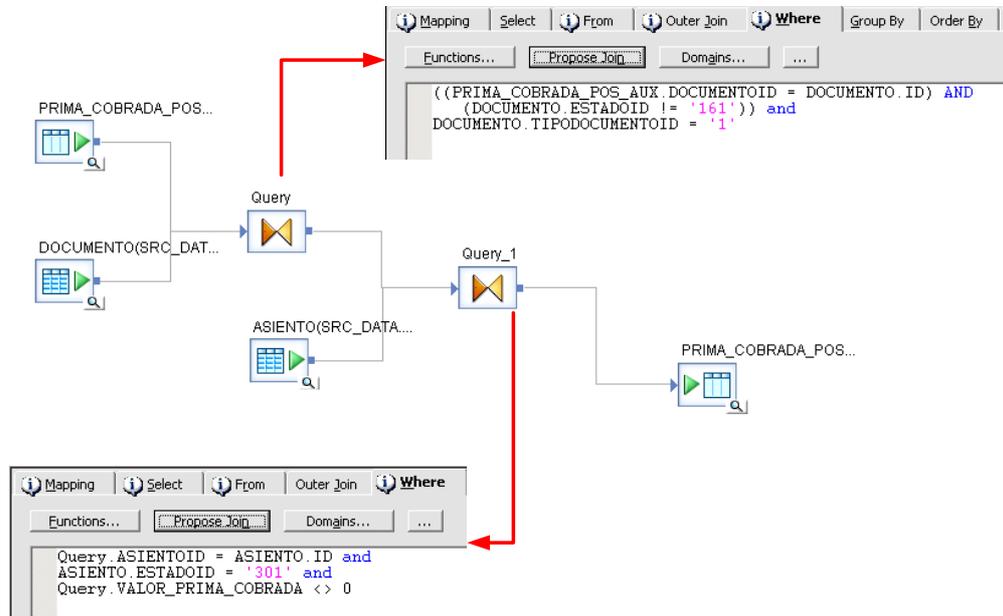


Figura 3.39 Flujo de datos DF_LOAD_PRIMA_COBRADA_002

El flujo de datos DF_LOAD_PRIMA_COBRADA_003 cruza las tablas fuentes que contienen las claves naturales de las dimensiones. Además, extrae aquellos endosos que se encuentren en su último estado y cuyas coberturas afecten al valor de primas y valor asegurado. Los datos extraídos son cargados en la tabla temporal PRIMA_COBRADA_NEG_AUX del esquema STAGE.

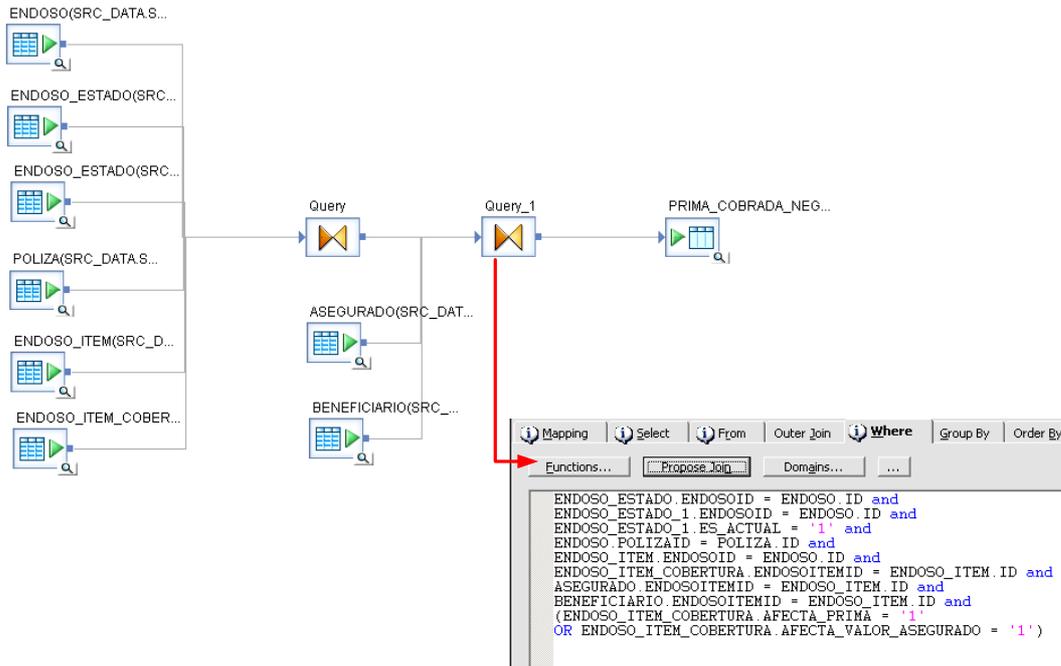


Figura 3.40 Flujo de datos DF_LOAD_PRIMA_COBRADA_003

El flujo de datos DF_LOAD_PRIMA_COBRADA_004 cruza la tabla temporal PRIMA_COBRADA_NEG_AUX con las tablas fuente DOCUMENTO y ASIENTO para obtener la producción cancelada a través de las notas de crédito que fueron contabilizadas como se muestra en la Figura 3.41.

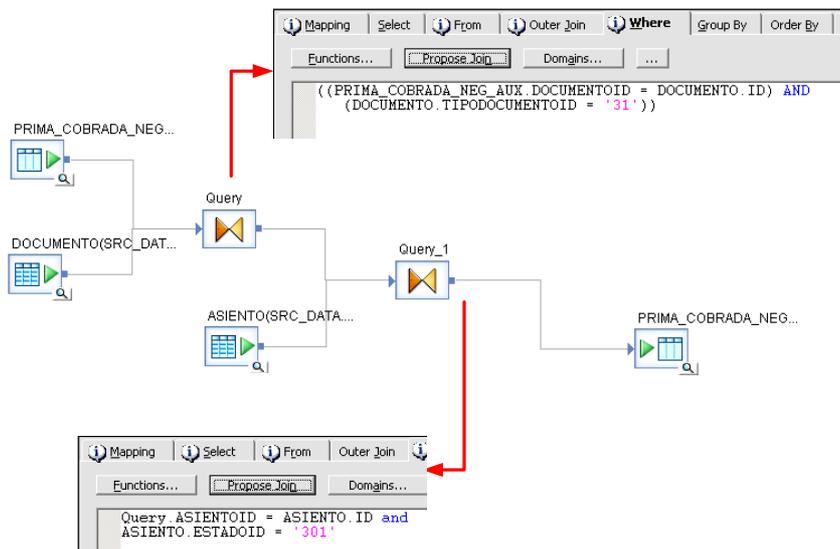


Figura 3.41 Flujo de datos DF_LOAD_PRIMA_COBRADA_004

Los flujos de datos DF_LOAD_PRIMA_COBRADA_005 y DF_LOAD_PRIMA_COBRADA_006 extraen el valor asegurado por extra asegurado. El flujo de datos DF_LOAD_PRIMA_COBRADA_005 cruza las tablas fuente del esquema DF_LOAD_PRIMA_COBRADA_001, adicionando la tabla fuente EXTRAS bajo las mismas condiciones. De igual forma el flujo de datos DF_LOAD_PRIMA_COBRADA_006 realiza un proceso similar que el flujo de datos DF_LOAD_PRIMA_COBRADA_002. Los datos extraídos por estos flujos de datos son cargados en las tablas temporales PRIMA_PAGADA_POS_EXT_AUX y PRIMA_COBRADA_POSITIVA_EXTRAS del esquema STAGE.

El flujo de datos DF_LOAD_PRIMA_COBRADA_007 realiza una unión de las tablas temporales PRIMA_COBRADA_POSITIVA, PRIMA_COBRADA_NEGATIVA y PRIMA_COBRADA_POSITIVA_EXTRAS del esquema STAGE. Además, también realiza una agrupación por las claves naturales de las dimensiones y los campos de fecha y realiza una agregación de los campos numéricos que serán cargados posteriormente en la tabla de hechos como se puede observar en la Figura 3.42.

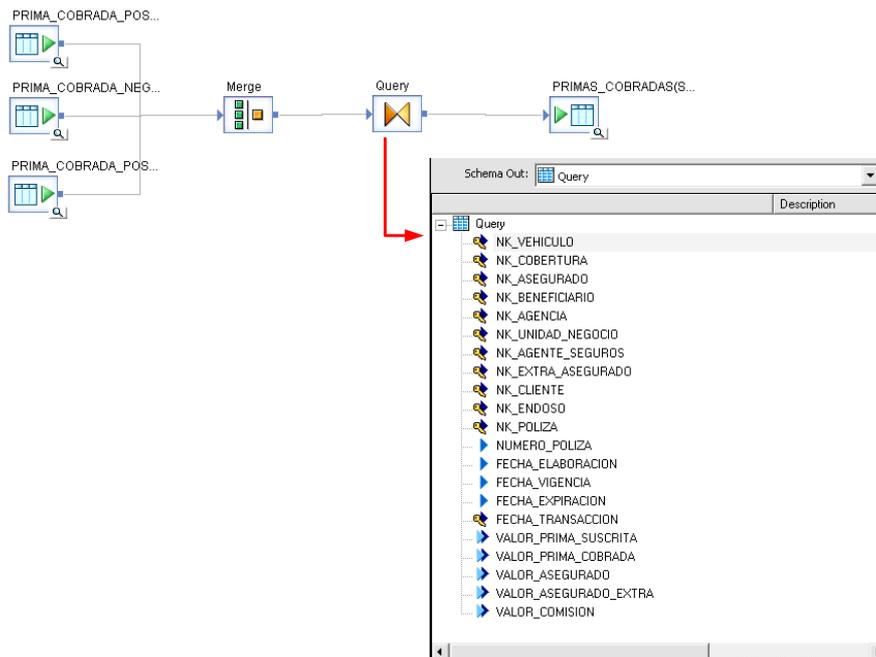


Figura 3.42 Flujo de datos DF_LOAD_PRIMA_COBRADA_007

ETL_SINIESTROS_POLIZAS

El flujo de trabajo ETL_SINIESTROS_POLIZAS se encarga de cargar los datos correspondientes a las indemnizaciones realizadas por reclamo y a las reservas económicas que generan los mismos. El flujo de trabajo contiene los flujos de trabajo ETL_SINIESTROS_BASE, ETL_INDEMNIZACIONES y ETL_RESERVAS como se puede observar en la Figura 3.43.



Figura 3.43 Estructura del flujo de trabajo ETL_SINIESTROS_POLIZAS

El flujo de trabajo ETL_SINIESTROS_BASE contiene los flujos de datos DF_LOAD_SINIESTROS_BASE_001 y DF_LOAD_SINIESTROS_BASE_002 encargados de cargar la tabla temporal SINIESTROS_BASE en el esquema STAGE. La tabla de siniestros base contiene todas las claves naturales y los campos no agregables que se cargarán en la tabla de hechos de siniestros FACT_SINIESTROS.

El flujo de datos DF_LOAD_SINIESTROS_BASE_001 cruza las tablas fuentes ENDOSO, ENDOSO_ITEM, RECLAMO y SUBRECLAMO y realiza siguientes tareas de limpieza: para los valores nulos del número de reclamo coloca 0 y para las fechas de apertura de reclamos nulas coloca la fecha 01/01/2005 como se puede observar en la Figura 3.44. Todos estos datos son cargados en la tabla temporal SINIESTROS_BASE_AUX del esquema STAGE.

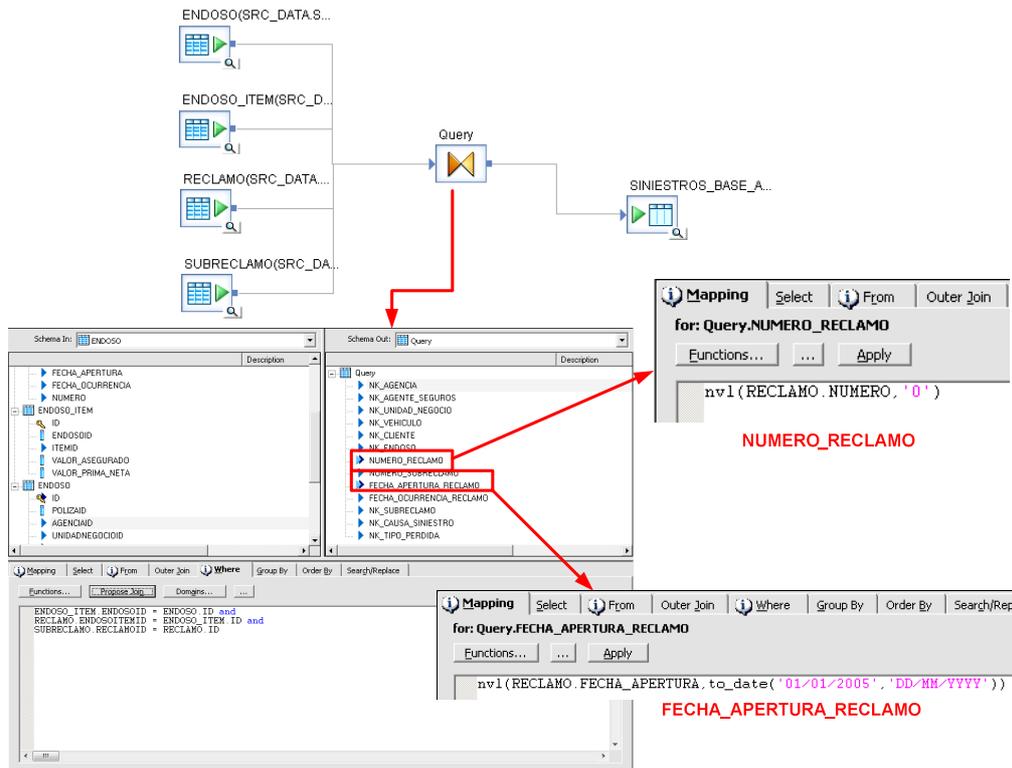


Figura 3.44 Flujo de datos DF_LOAD_SINIESTROS_BASE_001

El flujo de datos DF_LOAD_SINIESTROS_BASE_002 realiza una juntura con las tablas fuentes CAUSA_SINIESTRO y TIPO_PERDIDA. Se controló la integridad referencial realizando OUTER JOINS entre las tablas CAUSA_SINIESTRO y TIPO_PERDIDA y se cargó los datos en la tabla temporal SINIESTROS_BASE del esquema STAGE. La tabla siniestros base sirve como fuente para cargar los valores de indemnizaciones y los valores de reservas realizando las junturas con las tablas fuente necesarias como se explican en los flujos de trabajo ETL_INDEMNIZACIONES y ETL_RESERVAS.

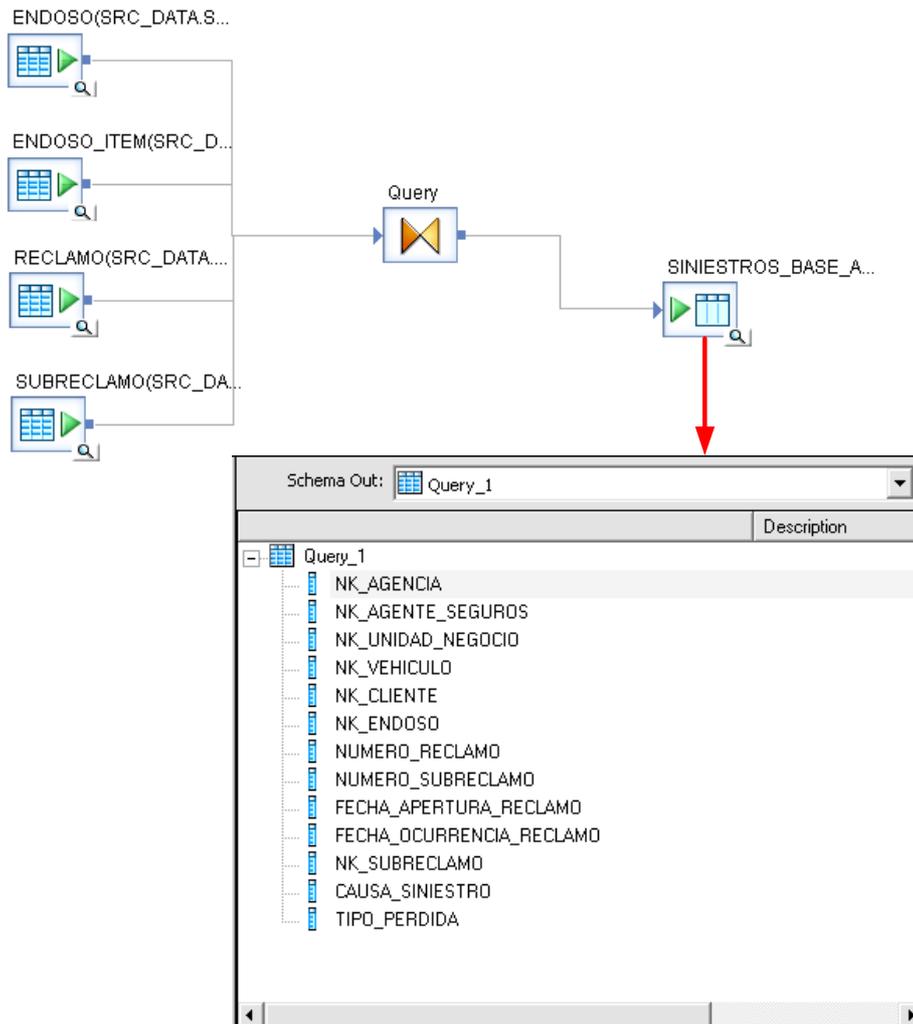


Figura 3.45 Flujo de datos DF_LOAD_SINIESTROS_BASE_002

El flujo de trabajo ETL_INDEMNIZACIONES se encarga de cargar los valores de indemnizaciones desglosados por valores de gastos, pagos, deducibles, anticipos y liquidaciones. Este flujo de trabajo contiene el flujo de datos DF_LOAD_INDEMNIZACIONES que cruza la tabla temporal SINIESTROS_BASE con la tabla fuente INDEMNIZACION. Además por cada tipo de indemnización agrega los valores de pagos, gastos, anticipos y liquidaciones agregados por cada clave natural que contendrá la tabla de hechos de siniestros, como se puede observar en

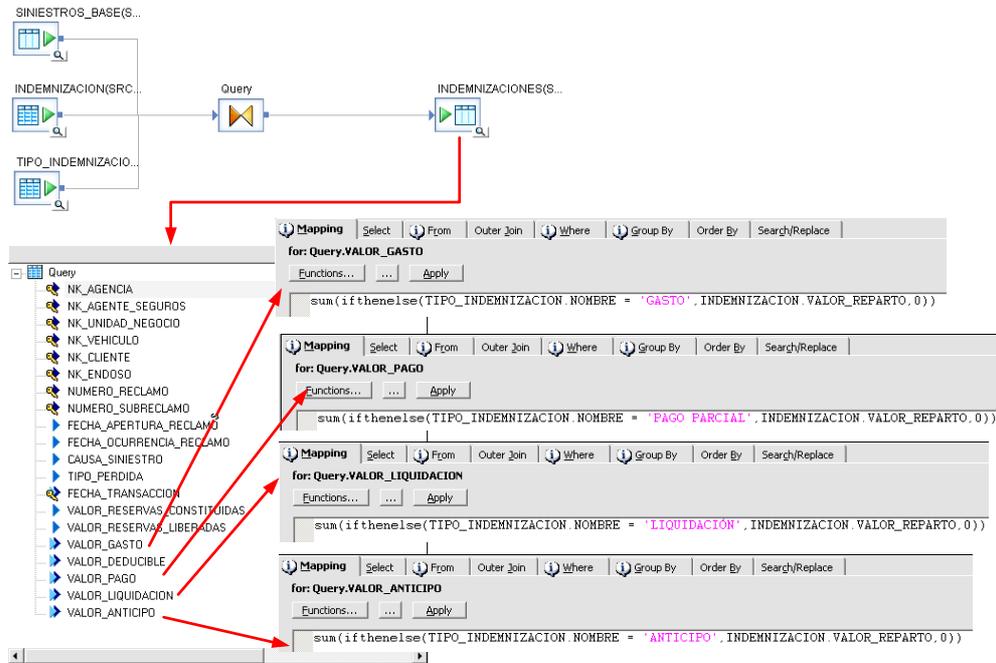


Figura 3.46 Flujo de datos DF_LOAD_INDEMNIZACIONES

El flujo de trabajo ETL_RESERVAS contiene el flujo de datos DF_LOAD_RESERVAS encargados de cargar el valor de reservas constituidas y liberadas. El flujo de datos DF_LOAD_RESERVAS cruza la tabla temporal SINIESTROS_BASE con la tabla fuente RESERVAS para agregar el valor de reservas constituidas y liberadas validando el tipo de reserva y agrupando estos valores por cada clave natural que será cargada en la tabla de hechos de siniestros como se muestra en la Figura 3.47. Los datos de este flujo son cargados en la tabla temporal RESERVAS del esquema STAGE.

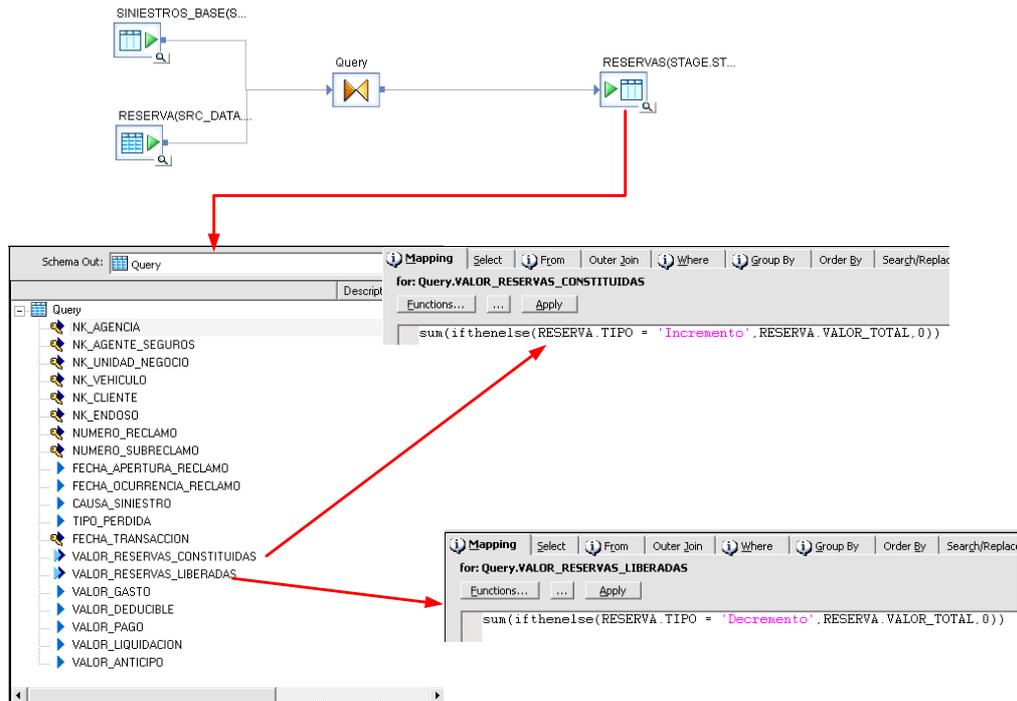


Figura 3.47 Flujo de datos DF_LOAD_RESERVAS

3.4.3.3.2 Flujo de Trabajo WF_STAGE_DIM_DATA

El flujo de trabajo WF_STAGE_DIM_DATA contiene flujos de datos que extraen los datos de las tablas temporales del esquema STAGE, aplican la técnica de control de cambios SCD Tipo 2 y cargan de datos a las dimensiones y tablas de hechos del esquema DIM_DATA.

ETL_DIM_COBERTURA

El flujo de trabajo ETL_DIM_COBERTURA contiene el flujo de datos DF_LOAD_DIM_COBERTURA y se encarga de cargar la dimensión cobertura (DIM_COBERTURA) del esquema DIM_DATA. Para cargar ésta dimensión se extrajeron los datos de la tabla temporal MAPEO_COBERTURA del esquema STAGE, cargándose la dimensión únicamente con valores de coberturas únicas y con claves subrogadas enteras como se puede observar en la Figura 3.48. En la carga de ésta dimensión no se aplicó la técnica de control de cambios SCD.

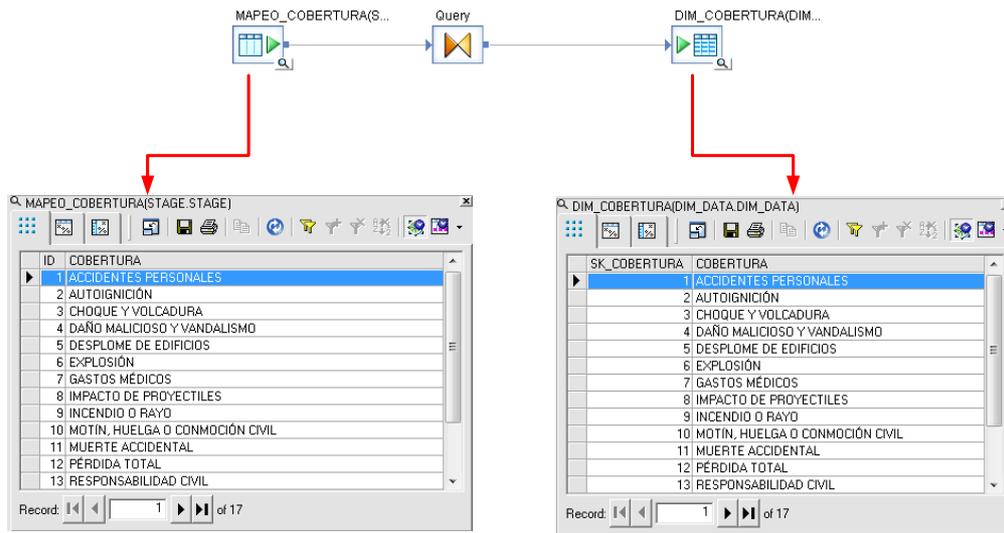


Figura 3.48 Flujo de datos DF_LOAD_DIM_COBERTURA

ETL_DIM_AGENCIA

El flujo de trabajo ETL_DIM_AGENCIA contiene el flujo de datos DF_LOAD_DIM_AGENCIA para cargar la dimensión AGENCIA (DIM_AGENCIA) del esquema DIM_DATA. En el flujo de datos para cargar ésta dimensión se aplicó la técnica de control de cambios SCD (Slowly Changing Dimension) Tipo 2 a través del uso de los siguientes transformadores del Diseñador de Data Integrator:

Transformador	Icono	Descripción
Table_Comparison		<p>Realiza una comparación de columnas registro por registro tomando como referencia la clave primaria de las tablas que se va a comparar. Existen dos métodos de comparación:</p> <ul style="list-style-type: none"> El método Cached Comparison Table es importante para la comparación de tablas con muchos registros. El método Row-by-row select es recomendado para realizar

		<p>comparaciones de tablas con pocos registros.</p> <p>La comparación se la realiza con la tabla destino a cargar, por lo que los campos del transformador Query deben ser los mismos que la tabla destino.</p>
History_Preserving		<p>Se encarga en base a la salida del transformador de comparación de evaluar los registros que han cambiado. Coloca una bandera con valor 1 sí el registro ha cambiado y 0 para preservar los registros anteriores. Además registra las fechas desde que son válidos y hasta cuando.</p>
Key_Generation		<p>Se encarga de generar un número secuencial incrementado por 1 para generar la clave subrogada de cada dimensión.</p>

Tabla 3.20 Transformadores del Diseñador de Data Integrator para aplicar la técnica de control de cambios SCD Tipo 2

En el flujo de datos DF_LOAD_DIM_AGENCIA se realiza la juntura entre la tabla temporal AGENCIA_ENCARGADO del esquema STAGE y la tabla fuente AGENCIA. La tabla AGENCIA_ENCARGADO contiene el nombre de los encargados de cada agencia. En la dimensión DIM_AGENCIA se crearon los campos ACTIVO, VALIDO_DESDE, VALIDO_HASTA y la clave natural NK_AGENCIA.

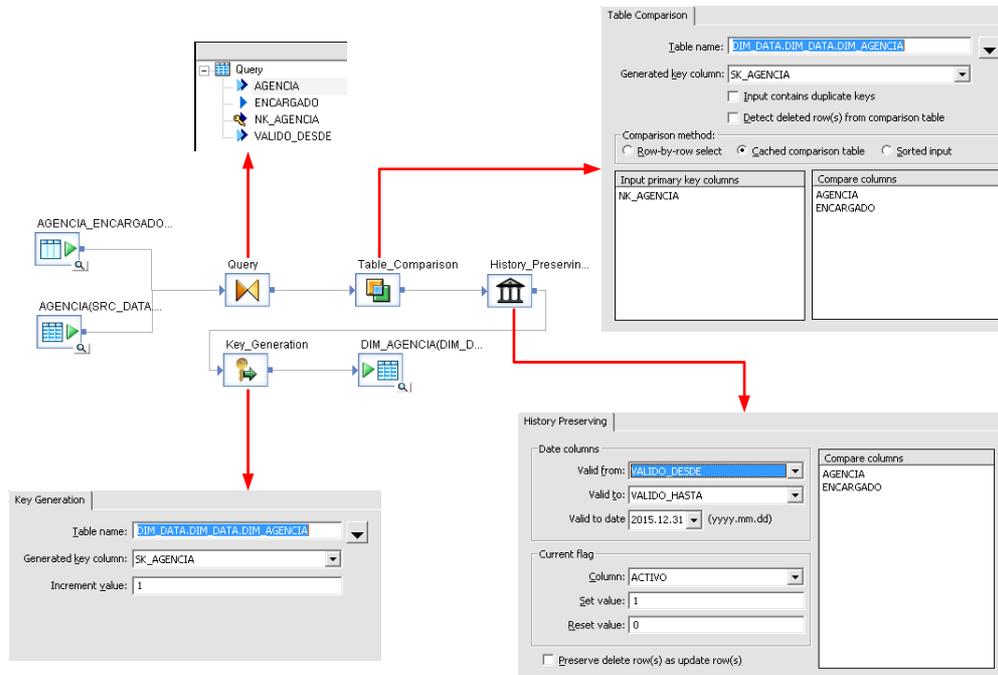


Figura 3.49 Flujo de datos DF_LOAD_DIM_AGENCIA

En la Figura 3.49 se puede observar el Query de donde se extrajeron los datos del nombre de agencia, encargado, la clave natural de la tabla fuente AGENCIA y se creó la columna VALIDO_DESDE con la función `sysdate()` para registrar la fecha del sistema en la cual se ha insertado el registro. En el transformador Table_Comparison se seleccionó la dimensión DIM_AGENCIA del esquema DIM_DATA, la columna de la clave subrogada (SK_AGENCIA), la clave primaria NK_AGENCIA y las columnas a comparar (AGENCIA y ENCARGADO). En el transformador History_Preserving se selecciona la columna VALIDO_DESDE en la opción Valid from, VALIDO_HASTA en la opción Valid to y en la opción Valid to date se seleccionó 2015.12.31. En la opción de Current Flag, se seleccionó la columna ACTIVO y las opciones Set value como 1 y Reset Value como 0. Además, en las columnas de comparación se arrastró las columnas AGENCIA y ENCARGADO. El transformador Key_Generation se encarga de generar la clave subrogada (SK_AGENCIA) de la dimensión DIM_AGENCIA incrementada en 1.

En resumen el flujo de datos comparará las columnas AGENCIA y ENCARGADO en base a la clave natural de la tabla AGENCIA (NK_AGENCIA). Se registrará la fecha en la cual el registro es insertado a través de la columna VALIDO_DESDE y

la fecha hasta la cual el registro es válido VALIDO_HASTA. Si los datos de encargado o nombre de agencia cambian se creará un nuevo registro con la columna ACTIVO = 1, se preservará el anterior registro con la columna ACTIVO = 0 y así para todos los cambios que existan en las columnas de nombre de agencia y encargado.

ETL_DIM_UNIDAD_NEGOCIO

El flujo de trabajo ETL_DIM_UNIDAD_NEGOCIO contiene el flujo de datos DF_LOAD_DIM_UNIDAD_NEGOCIO para cargar la dimensión DIM_UNIDAD_NEGOCIO del esquema DIM_DATA. De igual forma que en la dimensión Agencia (DIM_AGENCIA) se aplicó la técnica de control de cambios SCD Tipo 2, comparando la columna del nombre de la unidad de negocio como se puede observar en la Figura 3.50.

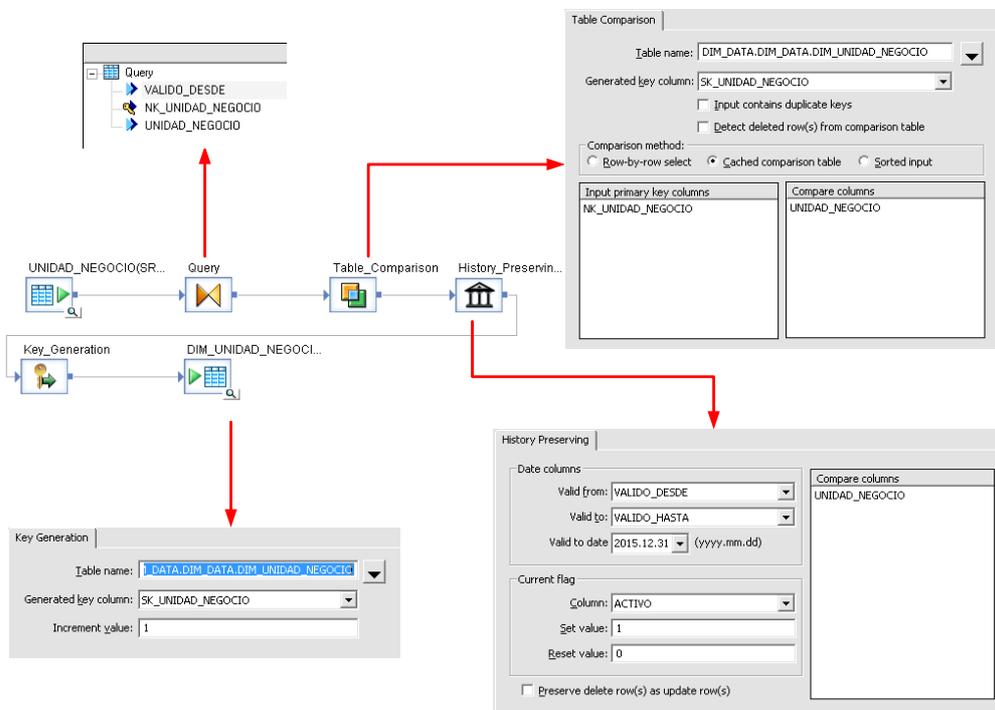


Figura 3.50 Flujo de datos DF_LOAD_DIM_UNIDAD_NEGOCIO

ETL_DIM_ASEGURADO

El flujo de trabajo ETL_DIM_ASEGURADO contiene el flujo de datos DF_LOAD_DIM_ASEGURADO para cargar la dimensión asegurado (DIM_ASEGURADO) del esquema DIM_DATA. Los datos de la tabla temporal MAPEO_ASEGURADO del esquema STAGE son transportados y cargados en la dimensión DIM_ASEGURADO manteniéndose únicamente un único registro del nombre del asegurado y con la clave subrogada generada por cada registro como se puede observar en la Figura 3.51.

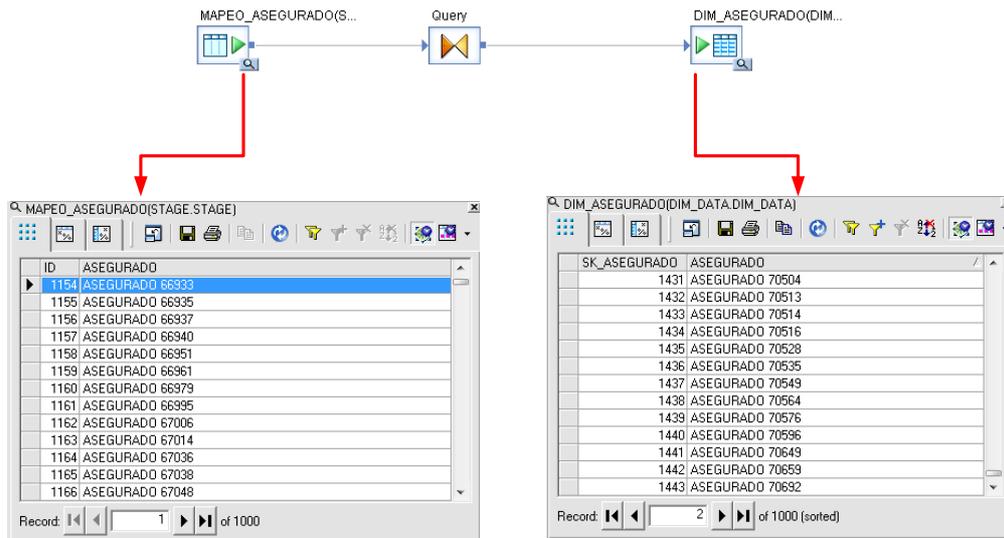


Figura 3.51 Flujo de datos DF_LOAD_DIM_ASEGURADO

ETL_DIM_BENEFICIARIO

El flujo de trabajo ETL_DIM_BENEFICIARIO contiene el flujo de datos DF_LOAD_DIM_BENEFICIARIO que se encarga de transportar los datos de la tabla temporal MAPEO_BENEFICIARIO del esquema STAGE a la dimensión beneficiario (DIM_BENEFICIARIO) del esquema DIM_DATA, como se muestra en la Figura 3.52.

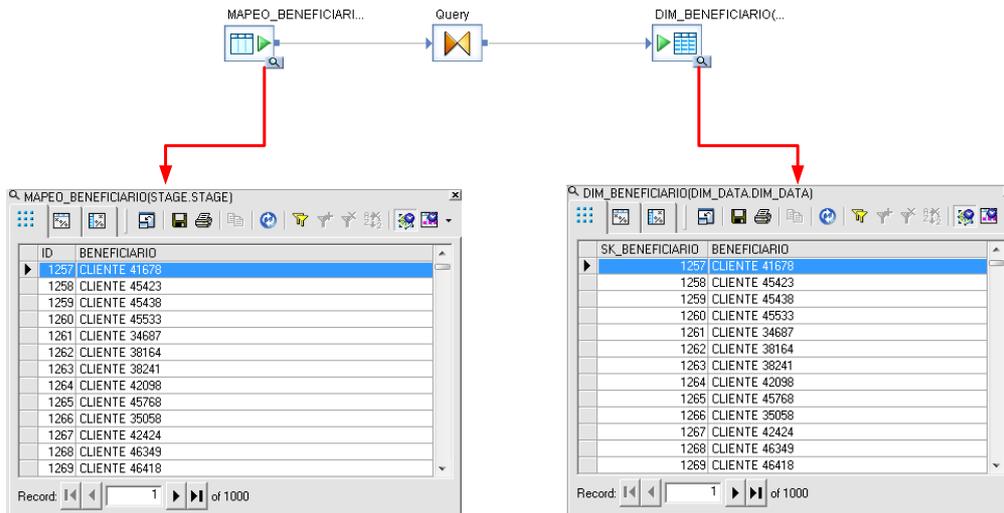


Figura 3.52 Flujo de datos DF_LOAD_DIM_BENEFICIARIO

ETL_DIM_EXTRA_ASEGURADO

El flujo de trabajo ETL_DIM_EXTRA_ASEGURADO contiene el flujo de datos DF_LOAD_DIM_EXTRA_ASEGURADO encargado de transportar los datos de la tabla temporal MAPEO_EXTRAS_ASEGURADOS del esquema STAGE a la dimensión DIM_EXTRA_ASEGURADO del esquema DIM_DATA, como se observa en la Figura 3.53.

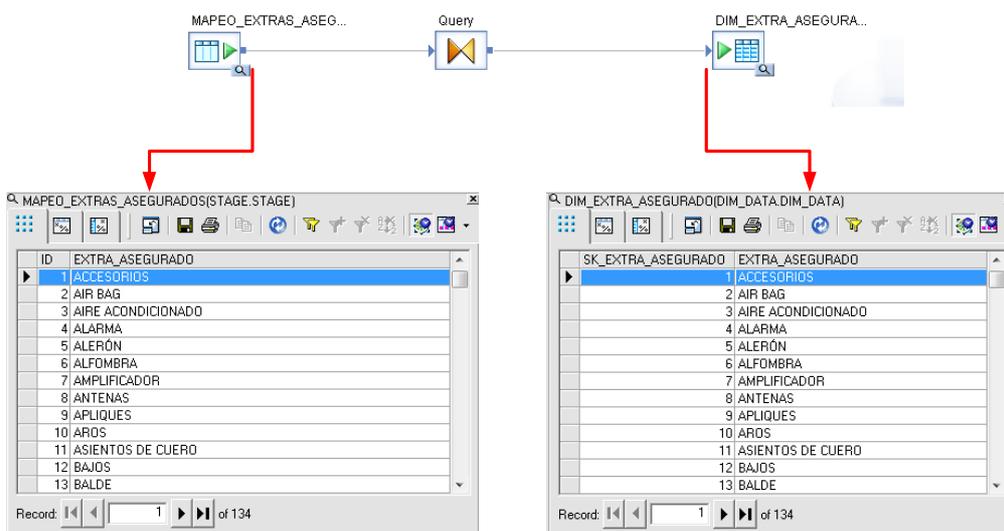


Figura 3.53 Flujo de datos DF_LOAD_DIM_EXTRA_ASEGURADO

ETL_DIM_VEHICULO

El flujo de trabajo ETL_DIM_VEHICULO contiene el flujo de datos DF_LOAD_DIM_VEHICULO el cual se encarga de transportar los datos de vehículos transformados de la tabla VEHICULO_AUX del esquema STAGE a la dimensión DIM_VEHICULO del esquema DIM_DATA. Para cargar la dimensión se aplicó la técnica de control de cambios SCD Tipo 2 comparando las columnas COLOR_VEHICULO, MARCA_VEHICULO, MODELO_VEHICULO, CLASE_VEHICULO, TIPO_VEHICULO, ZONA_RIESGO, PLACAS, MOTOR, CHASIS, ANIO_FABRICACION y NUMERO_OCUPANTES como se muestra en la Figura 3.54.

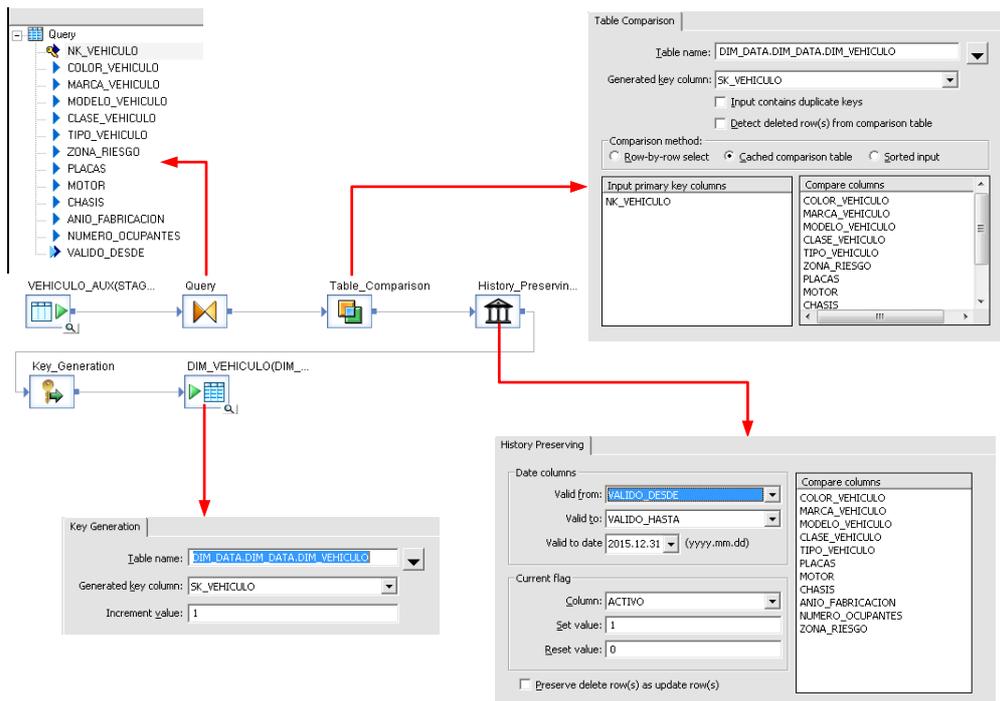


Figura 3.54 Flujo de datos DF_LOAD_DIM_VEHICULO

ETL_DIM_AGENTE_SEGUROS

El flujo de datos ETL_DIM_AGENTE_SEGUROS contiene el flujo de datos DF_LOAD_DIM_AGENTE encargado de transportar los datos de la tabla fuente AGENTE a la dimensión DIM_AGENTE_SEGUROS del esquema DIM_DATA. Para cargar la dimensión se aplicó la técnica de control de cambios SCD Tipo 2 comparando las columnas AGENTE_SEGUROS y ANIO_INICIO_OPERACION como se muestra en la Figura 3.55.

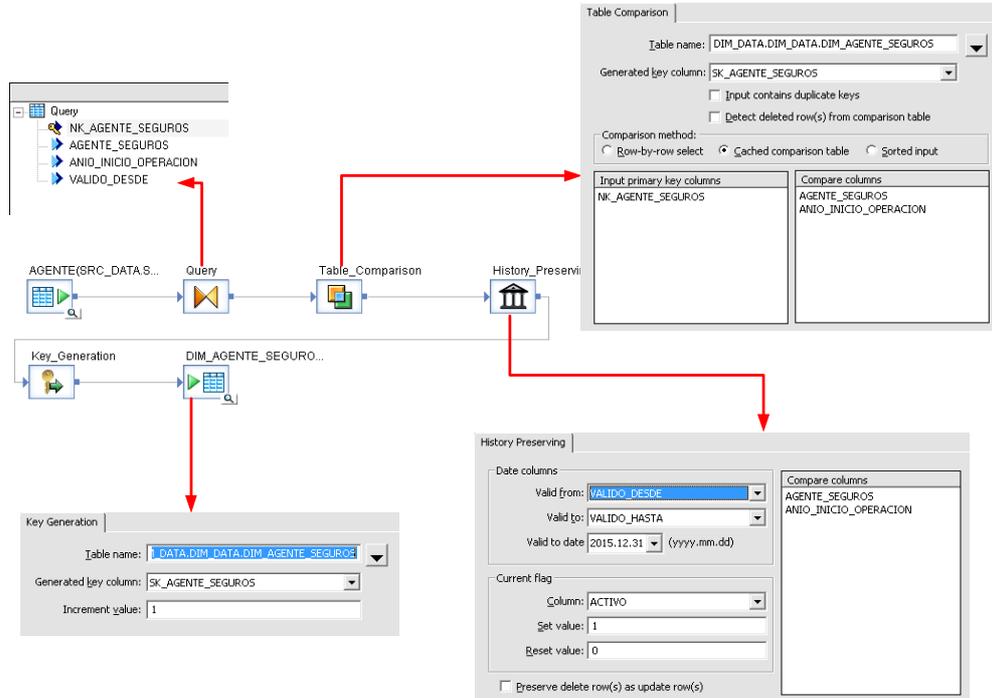


Figura 3.55 Flujo de datos DF_LOAD_DIM_AGENTE

ETL_DIM_CLIENTE

El flujo de trabajo ETL_DIM_CLIENTE contiene el flujo de datos DF_LOAD_DIM_CLIENTE para transportar los datos de clientes transformados de la tabla temporal CLIENTE_AUX del esquema STAGE a la dimensión DIM_CLIENTE del esquema DIM_DATA. En el flujo de datos también se aplica la técnica de control de cambios SCD Tipo 2 comparando las columnas CLIENTE y TIPO_CLIENTE como se muestra en la Figura 3.56.

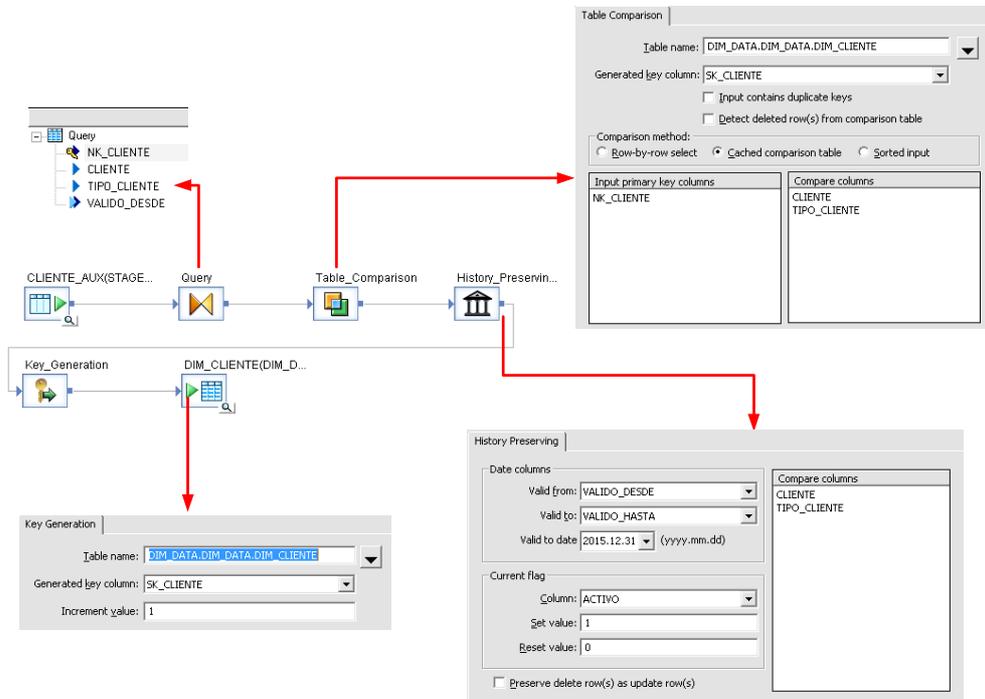


Figura 3.56 Flujo de datos DF_LOAD_DIM_CLIENTE

ETL_DIM_ENDOSO

El flujo de trabajo ETL_DIM_ENDOSO contiene el flujo de datos DF_LOAD_DIM_ENDOSO para transportar los datos transformados de endosos que se encuentran en la tabla temporal ENDOSO_AUX del esquema STAGE a la dimensión DIM_ENDOSO del esquema DIM_DATA. Para cargar ésta dimensión se aplicó la técnica de control de cambios SCD Tipo 2 comparando las columnas NUMERO_ENDOSO, TIPO_ENDOSO, ESTADO_ENDOSO, NUMERO_DOCUMENTO y TIPO_DOCUMENTO como se puede observar en la Figura 3.57.

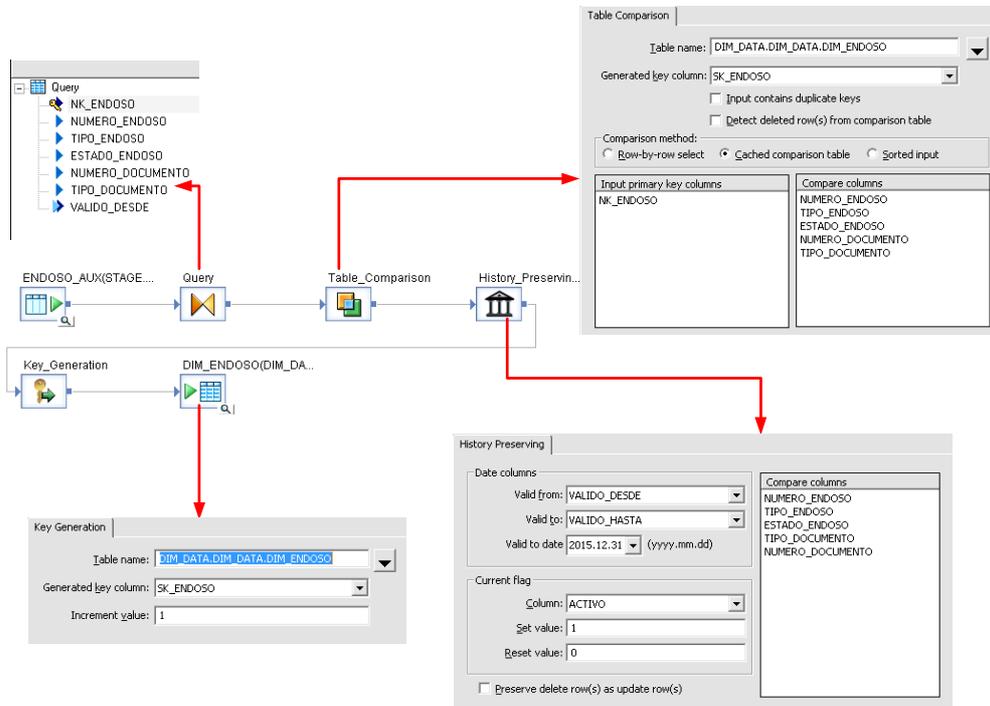


Figura 3.57 Flujo de datos DF_LOAD_DIM_ENDOSO

ETL_FACT_PRODUCCION

El flujo de trabajo ETL_FACT_PRODUCCION contiene los flujos de datos DF_LOAD_FACT_PRODUCCION_001, DF_LOAD_FACT_PRODUCCION_002, DF_LOAD_FACT_PRODUCCION_003 y DF_LOAD_FACT_PRODUCCION para cargar la tabla de hechos FACT_PRODUCCION del esquema DIM_DATA.

El flujo de datos DF_LOAD_FACT_PRODUCCION_001 realiza una unión de las tablas temporales PRIMAS_SUSCRITAS y PRIMAS_COBRADAS del esquema STAGE para cargar la tabla temporal FACT_PRODUCCION_AUX_01 del esquema STAGE. Se agregaron los valores de prima suscrita, prima cobrada, valor asegurado y valor comisión para las claves naturales y campos no agregables que serán cargados posteriormente en la tabla de hechos como se muestra en la Figura 3.58.

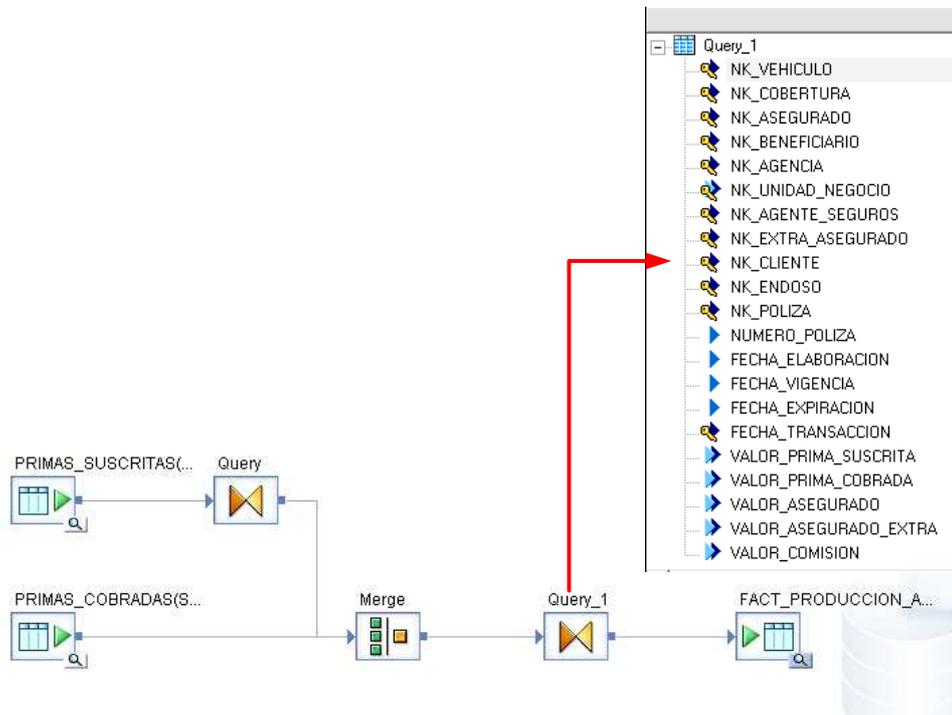


Figura 3.58 Flujo de datos DF_LOAD_FACT_PRODUCCION_001

El flujo de datos FACT_PRODUCCION_AUX_01 realiza una juntura entre la tabla temporal FACT_PRODUCCION_AUX_01 y las tablas de mapeo MAPEO_COBERTURA_AUX, MAPEO_ASEGURADO_AUX, MAPEO_BENEFICIARIO_AUX para obtener las claves naturales de las dimensiones. Además, se realiza la juntura con las dimensiones DIM_VEHICULO, DIM_AGENCIA, DIM_UNIDAD_NEGOCIO, DIM_AGENTE_SEGUROS, DIM_CLIENTE y DIM_ENDOSO para cargar los datos en la tabla temporal FACT_PRODUCCION_AUX_02 del esquema STAGE como se puede observar en la Figura 3.59. Es importante mencionar que en el WHERE del Query sólo se cruzan los valores que estén activos de las dimensiones filtrando la columna ACTIVO = 1.

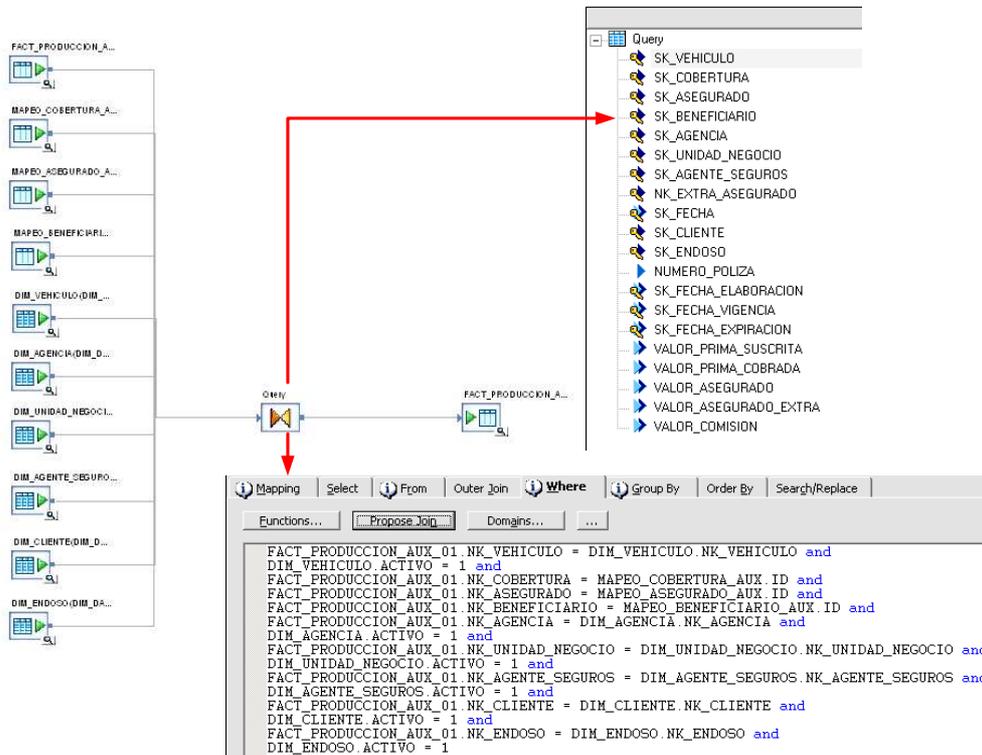


Figura 3.59 Flujo de datos DF_LOAD_FACT_PRODUCCION_002

El flujo de datos DF_LOAD_FACT_PRODUCCION_003 junta la tabla temporal FACT_PRODUCCION_AUX_02 del esquema STAGE con la tabla MAPEO_EXTRAS_ASEGURADOS_AUX y la dimensión DIM_FECHA cargando los datos en la tabla temporal FACT_PRODUCCION del esquema STAGE. En el Query se realizó un OUTER JOIN para extraer todos los datos ya que no todos los ítems tenían extras asegurados. El flujo de datos se puede observar en la Figura 3.60.

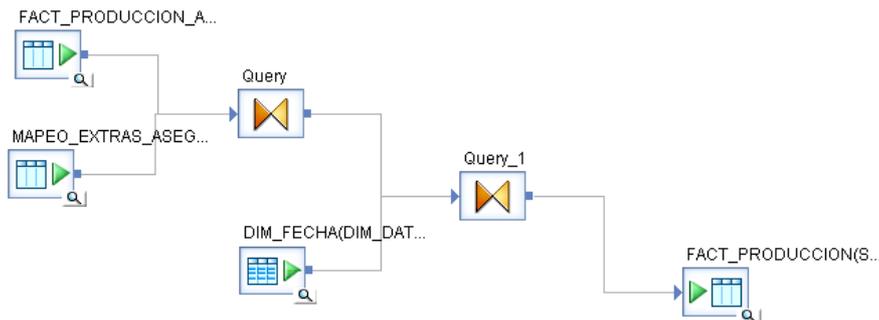


Figura 3.60 Flujo de datos DF_LOAD_FACT_PRODUCCION_003

El flujo de datos DF_LOAD_FACT_PRODUCION se encarga de transportar los datos de la tabla temporal FACT_PRODUCION del esquema STAGE a la tabla de hechos FACT_PRODUCION del esquema DIM_DATA, como se muestra en la Figura 3.61.

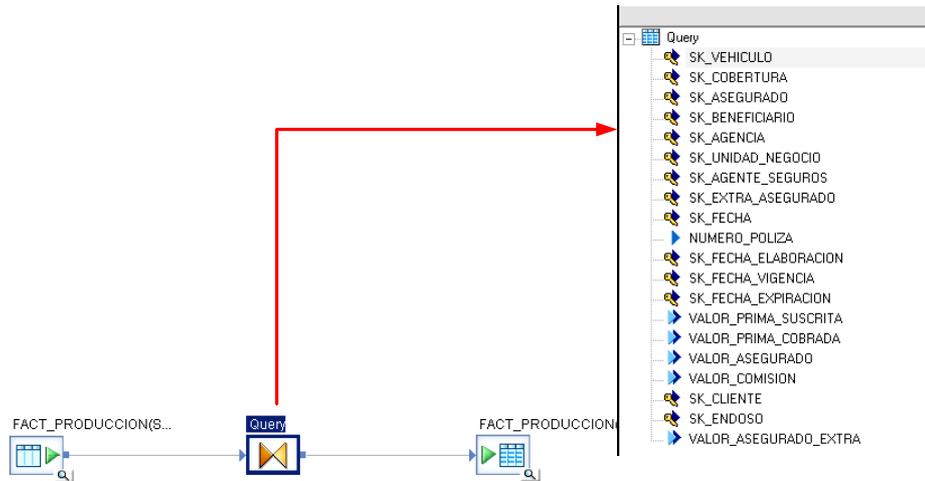


Figura 3.61 Flujo de datos DF_LOAD_FACT_PRODUCION

ETL_FACT_SINIESTROS

El flujo de trabajo ETL_FACT_SINIESTROS contiene los flujos de datos DF_LOAD_FACT_SINIESTROS_001, DF_LOAD_FACT_SINIESTROS_002 y DF_LOAD_FACT_SINIESTROS para cargar la tabla de hechos FACT_SINIESTROS del esquema DIM_DATA.

El flujo de datos DF_LOAD_FACT_SINIESTROS_001 realiza una unión de las tablas temporales RESERVAS e INDEMNIZACIONES del esquema STAGE y agrega los valores de reservas constituidas, reservas liberadas, gastos, pagos, deducibles, liquidaciones y anticipos para cargarlos en la tabla temporal FACT_SINIESTROS_AUX del esquema STAGE como se muestra en la Figura 3.62.

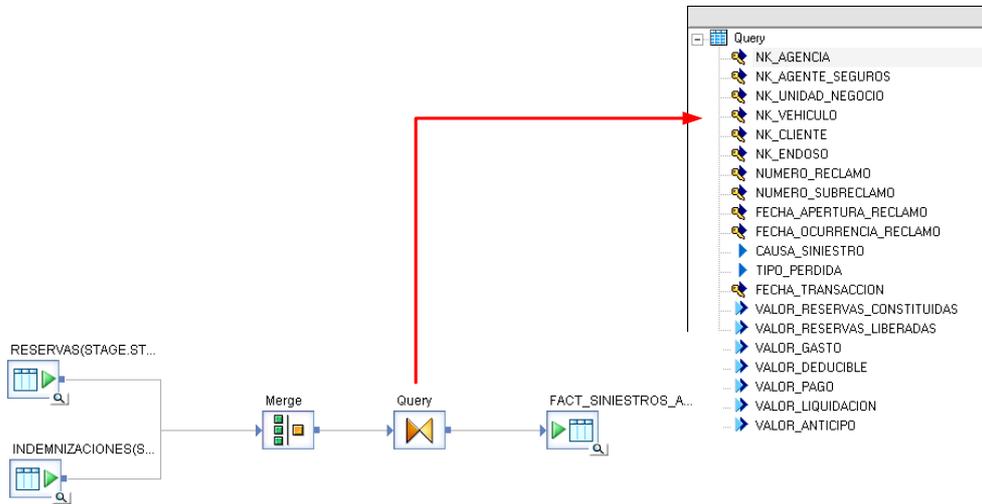


Figura 3.62 Flujo de datos DF_LOAD_FACT_SINIESTROS_001

El flujo de datos DF_LOAD_FACT_SINIESTROS_002 cruza la tabla temporal FACT_SINIESTROS_AUX cargada con las dimensiones DIM_CLIENTE, DIM_ENDOSO, DIM_AGENCIA, DIM_AGENTE_SEGUROS, DIM_UNIDAD_NEGOCIO y DIM_VEHICULO. Además, en el WHERE del Query sólo se cruzan las dimensiones activas filtrando el campo ACTIVO = 1, como se muestra en la Figura 3.63. Los datos son cargados en la tabla temporal FACT_SINIESTROS del esquema STAGE.

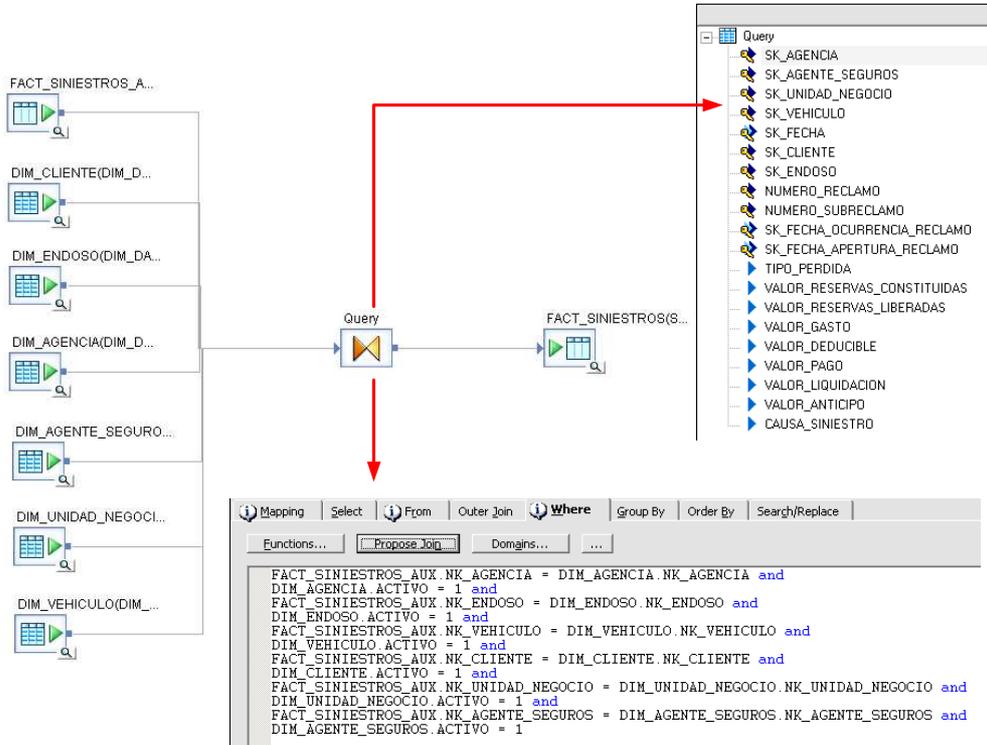


Figura 3.63 Flujo de datos DF_LOAD_FACT_SINIESTROS_002

El flujo de datos DF_LOAD_FACT_SINIESTROS carga la tabla de hechos FACT_SINIESTROS. Cruza la dimensión DIM_FECHA con la tabla temporal FACT_SINIESTROS como se puede observar en la Figura 3.64.

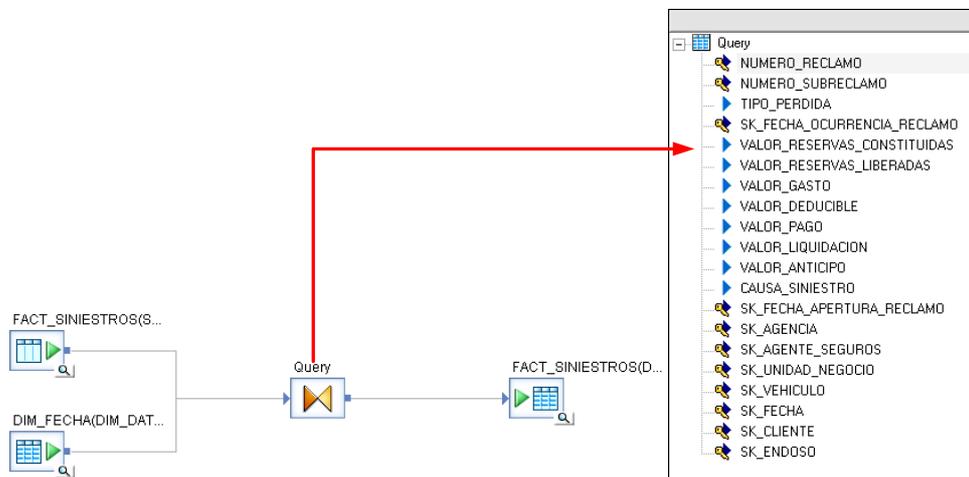


Figura 3.64 Flujo de datos DF_LOAD_FACT_SINIESTROS

ETL_FACT_PRODUCCION_SINIESTROS_MES

El flujo de trabajo ETL_FACT_PRODUCCION_SINIESTROS_MES contiene el flujo de datos DF_LOAD_PRODUCCION_SINIESTROS que se encarga de cargar la tabla de hechos FACT_PRODUCCION_SINIESTROS_MES, ésta tabla contiene los indicadores de prima cobrada, valor asegurado, comisiones, reservas, pagos y gastos y será utilizada como tabla de hechos agregada para consumir sus datos a través de un tablero de mando.

El flujo de datos DF_LOAD_PRODUCCION_SINIESTROS contiene un transformador SQL, que contiene la sentencia SQL que será encargada de cargar los datos de la tabla de hechos agregada. Además, los datos que se cargarán en la tabla de hechos DF_LOAD_PRODUCCION_SINIESTROS estarán agregados para cada mes colocando como fecha el primer día de cada mes.

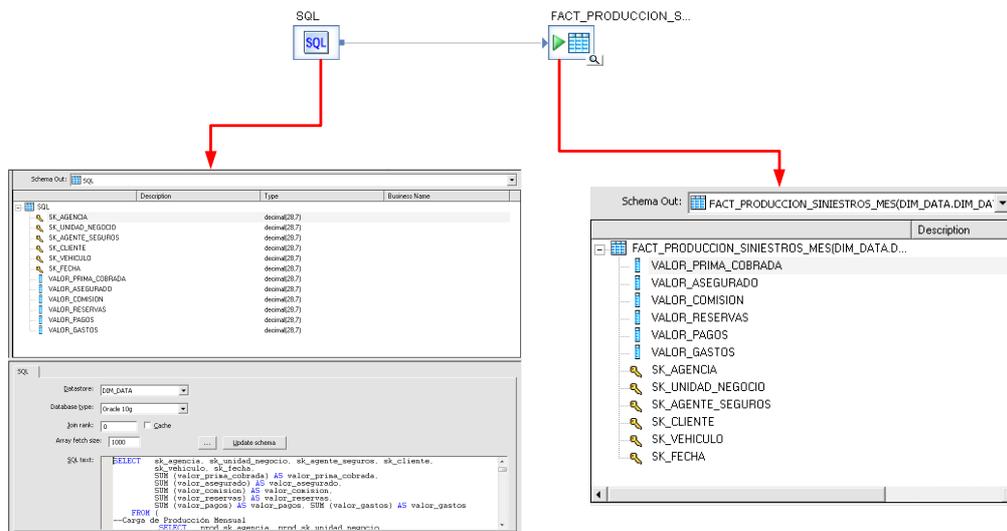


Figura 3.65 Flujo de datos DF_LOAD_PRODUCCION_SINIESTROS

En la Figura 3.65 se puede observar el transformador SQL y cómo éste carga la tabla de hechos FACT_PRODUCCION_SINIESTROS_MES. La sentencia SQL del transformador consume datos de las tablas de hechos FACT_PRODUCCION y FACT_SINIESTROS, los une y los agrega mensualmente por agencia, unidad de negocio, agente de seguros, cliente, vehículo y fecha.

3.4.4 CREACIÓN DEL UNIVERSO DE DATOS

Una vez construido el Data Mart para análisis de seguros de vehículos (DM_SEGUROS_VEHICULOS) es necesario explotar los datos a través de actividades análisis. Las actividades de análisis pueden ser realizadas a través de reportes o tableros de mando.

La fuente de datos para crear tableros o reportes dentro de la plataforma de inteligencia de negocios BusinessObjects Enterprise XI se conoce como universo. Un universo se construye usando la herramienta BusinessObjects Designer. Las tablas destino de los flujos de datos que cargan el Data Mart a través de Data Integrator son integrados con un universo creado en el Diseñador de BusinessObjects lográndose obtener metadatos sobre el linaje de datos, impacto en reportes y tableros, e integrando todos los procesos del sistema de Data Warehouse a través de la estandarización de la Arquitectura de Inteligencia de Negocios de una Compañía.

Los pasos de instalación de la plataforma BusinessObjects Enterprise XI Release 2 Service Pack 3 se describen en el ANEXO D. Además, en el ANEXO D se explica como crear la conexión de base de datos que será usada por cada universo que se defina.

Comentario [MM5]: Incluir el número de anexo que explica paso a paso la instalación de BusinessObjects Enterprise XI.

Comentario [MM6]: Incluir el número de anexo que explica paso a paso la instalación de BusinessObjects Enterprise XI.

Desde el Diseñador de Data Integrator se creó el universo de datos a través de la opción Tools ► BusinessObjects Universes ► Create... Aquí se seleccionó el almacén de datos que será fuente para crear el universo. Para el caso del proyecto, el Datastore DIM_DATA como se muestra en la Figura 3.66.

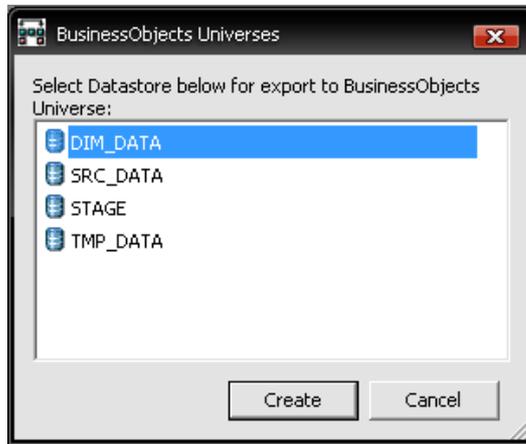


Figura 3.66 Selección del Datastore para crear un universo

Después de hacer clic en el botón Create y acceder a la herramienta Business Objects Designer, se utilizó el asistente para la generación de universos de datos como se muestra en la Figura 3.67.



Figura 3.67 Asistente para la generación de universos de datos

Se seleccionó todos los elementos del Datastore DIM_DATA que constituyen las dimensiones y tablas de hechos del Data Mart como se puede observar en la Figura 3.68.



Figura 3.68 Selección de elementos para crear el universo

Después de hacer clic en el botón Next, se seleccionó la conexión DWH y se colocó el nombre de universo como DM_VEHIC como se muestra en la Figura 3.69.

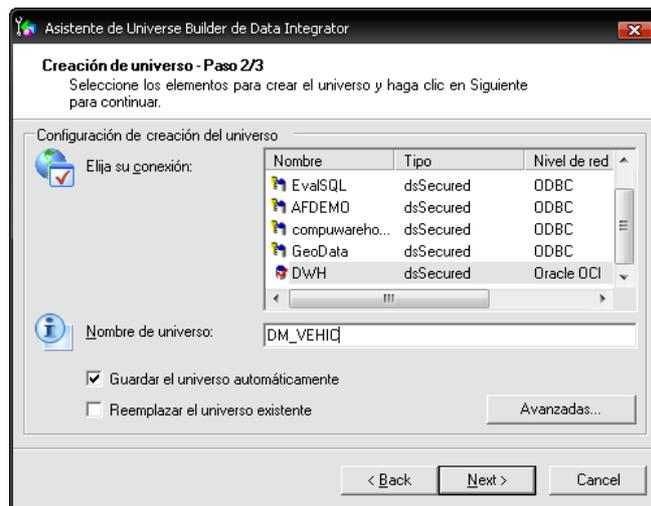


Figura 3.69 Configuración de conexión y nombre del universo

Después de hacer clic en el botón Next se mostró el resumen de configuración del universo, creándose el universo enlazado con los flujos de carga de datos de Data Integrator como se muestra en la Figura 3.70.

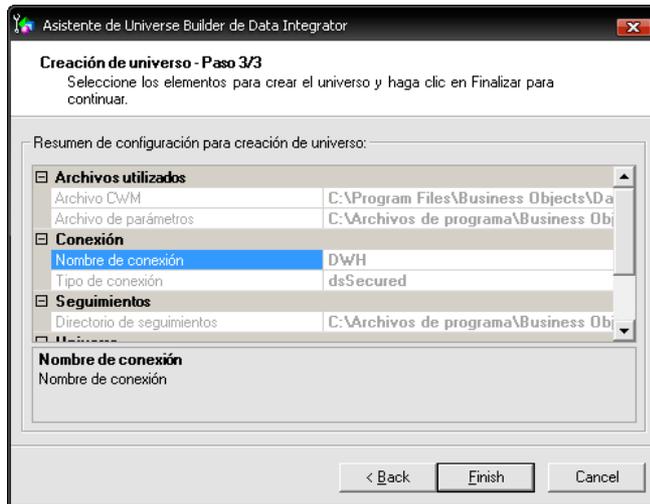


Figura 3.70 Resumen de configuración de creación de universos

Después de hacer clic en el botón Finish, se abrió el Diseñador de BusinessObjects con el universo de datos DM_VEHIC creado como se muestra en la Figura 3.71.

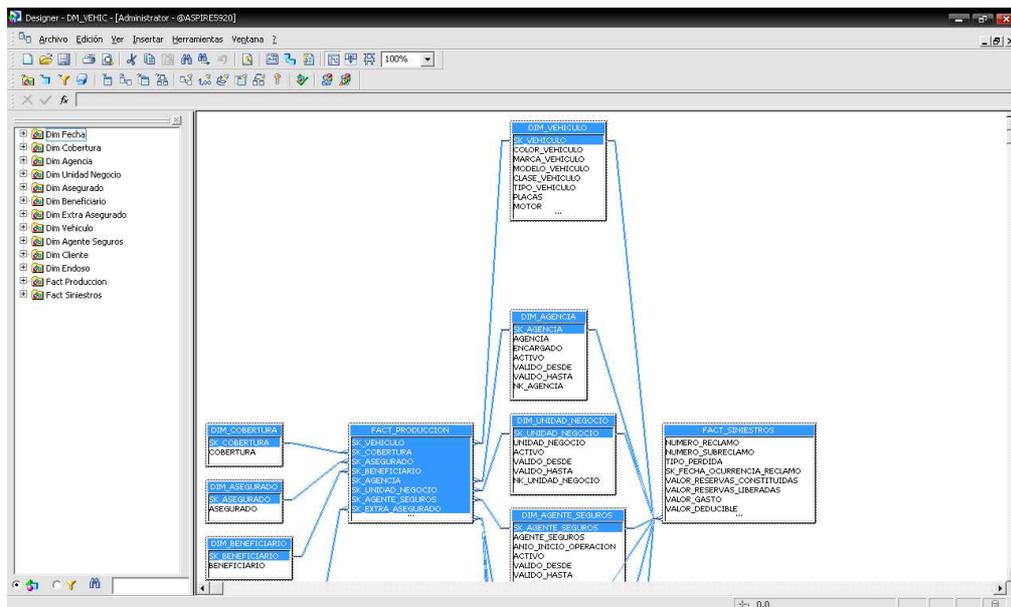


Figura 3.71 Universo de datos DM_VEHIC

Una vez creado el universo se modificaron los nombres de clases, dimensiones e indicadores y se realizaron las operaciones necesarias para que el universo cumpla con los requerimientos funcionales. En el ANEXO E se describe los metadatos del universo DM_VEHIC.

Comentario [MM7]: Anexo con metadatos del universo para reportes DM_VEHIC.

Posterior a la creación del universo DM_VEHIC, se creó el universo DM_VEHIC_PM que será usado por la herramienta de Gestión del Rendimiento y permitirá crear las métricas y analíticas que serán usadas en el tablero de mando.

El universo DM_VEHIC_PM utiliza la conexión DWH creada anteriormente y contiene las dimensiones Agencia, Encargado de Agencia, Unidad Negocio, Agente de Seguros, Marca Vehículo, Modelo Vehículo y los indicadores de valor asegurado, valor prima cobrada, valor comisión, valor siniestros pagados, valor reservas, resultado técnico, tasa promedio, porcentaje de comisiones y porcentaje de siniestralidad como se muestra en la Figura 3.72.

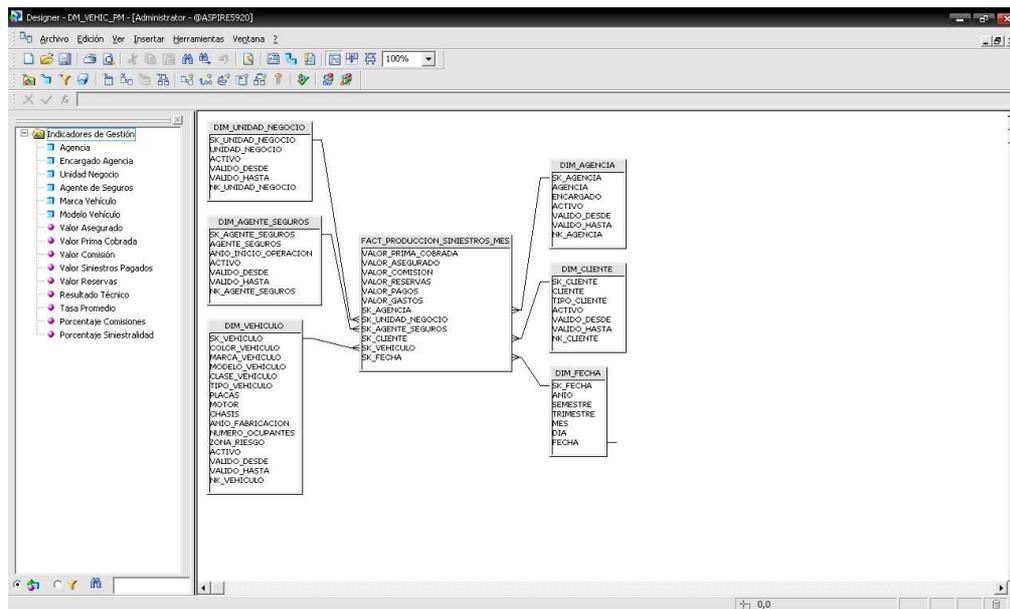


Figura 3.72 Universo de datos DM_VEHIC_PM

El universo de datos DM_VEHIC_PM contiene un auto contexto en la dimensión fecha, el auto contexto sobre ésta dimensión es obligatorio para crear tableros usando la herramienta de Gestión del Rendimiento de BusinessObjects. En la Figura 3.73 se puede observar la dimensión DIM_FECHA y el auto contexto creado.

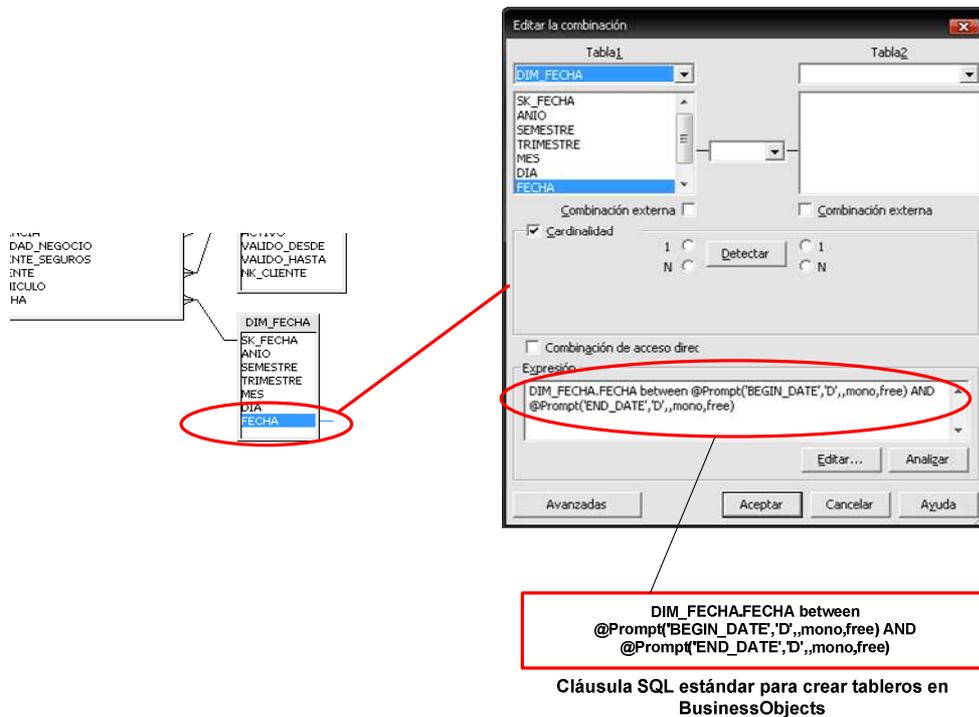


Figura 3.73 Auto contexto sobre la dimensión fecha para crear tableros de mando en BusinessObjects Performance Management

En el **ANEXO F** se describe los metadatos del universo DM_VEHIC_PM y las fórmulas con las cuáles se crearon los indicadores de análisis del universo. Es importante volver a mencionar que éste universo sirve como fuente para mostrar las métricas y analíticas en el tablero de control creado en la herramienta BusinessObjects Performance **Management**.

Comentario [MM8]: Anexo con metadata del universo del tablero DM_VEHIC_PM.

Comentario [MM9]: Incluir una tabla con las dimensiones e indicadores de análisis de cada universo creado, así como también de los indicadores o medidas calculadas.

3.4.5 PRUEBAS

Después de haber implementado los procesos ETL para cargar el Data Mart de seguros de vehículos es necesario e imprescindible realizar las pruebas necesarias para asegurarse que la información cargada en el Data Mart sea consistente con la información de los reportes del sistema OLTP.

Los sistemas OLTP generalmente tienen reportes predefinidos que los usuarios pueden obtener ya sea desde opciones del sistema o solicitándolos al Departamento de Tecnologías de la Información de una Compañía. La desventaja principal de obtener reportes a través de cualquiera de las dos opciones

mencionadas anteriormente es que los reportes están predefinidos, es decir los usuarios no pueden generar reportes a demanda para responder preguntas del negocio. Las herramientas de inteligencia de negocios brindan a los usuarios finales éste tipo de facilidad con el objetivo de que puedan realizar el análisis de la información obtenida y puedan tomar las decisiones del negocio adecuadas en el momento preciso.

Para la realización de pruebas de carga de datos se usó dos reportes del sistema OLTP. Los reportes Detalle Producción y Detalle Siniestros elaborados en Crystal Reports, los mismos que consumen datos del esquema SRC_DATA que contiene el modelo relacional de seguros de vehículos. El reporte de Detalle Producción permitirá verificar que la carga de los indicadores de Producción se haya cargado correctamente en el Data Mart analizando un mes al azar. El reporte de Detalle Siniestros permitirá verificar que la carga de los indicadores de siniestros se haya cargado sin problemas, de igual forma analizando un mes al azar.

3.4.5.1 Cuadre de datos de producción

El reporte Detalle Producción muestra la producción de pólizas para un rango de fechas. El reporte muestra la marca, modelo, número de póliza, valor asegurado, prima cobrada neta y comisión pagada agrupada por cada unidad de negocio y agencia como se muestra en la Figura 3.74.

MARCA	MODELO	NRO. POLIZA	VALOR ASEGURADO	PRIMA NETA	COMISIÓN
AGENCIA CENTRO 1			476,872.20	12,456.93	1,742.39
NEGOCIOS INDIVIDUALES			415,693.51	11,824.75	1,650.82
NISSAN	MURANO-AT	1016561	54,630.00	815.33	122.29
CHEVROLET	OPTRA 1. 8L T.M.	202409	16,455.00	201.24	30.19
HONDA	CR-V-RD6843P-KK	1007622	24,000.00	368.00	55.19
KENWORTH	T800 6X4	1007726	100,845.00	3,630.42	544.56
HYUNDAI	SANTA-FE	1016995	22,500.00	320.00	0.00
CHEVROLET	OPTRA 1. 8L T.M.	1016382	17,794.00	190.12	28.51
FORD	ECOSPORT-XLT-4X2	1040	20,709.51	807.40	113.04
NISSAN	MURANO-AT	1013818	54,990.00	2,074.55	311.18
FORD	ECOSPORT-XLT-4X2	1036	21,350.00	832.39	116.53
CITROEN	C3 SX 14I	996	13,700.00	533.97	74.76
HYUNDAI	TUCSON GL	995	20,990.00	818.34	114.57
MITSUBISHI	CANTER	1074	21,100.00	822.31	115.12
HYUNDAI	SANTA-FE	1016995	20,150.00	244.80	0.00
HYUNDAI	ACCENT	1007622	6,480.00	165.88	24.88

Figura 3.74 Reporte de Detalle Producción Enero 2005 Primera Página

El reporte de Detalle Producción muestra seccionado los resultados por Agencia y Unidad de Negocio, además muestra agregados los datos de valor asegurado, prima cobrada neta y comisión. Además, muestra un gran total al final correspondiente a toda la producción generada para el rango de fechas ingresado como se muestra en la Figura 3.75.

MARCA	MODELO	NRO. POLIZA	VALOR ASEGURADO	PRIMA NETA	COMISION
AGENCIA VALLES 1			90,396.86	3,682.45	627.69
NEGOCIOS MASIVOS			90,396.86	3,682.45	627.69
CHEVROLET	OPTRA 1. 8L T.M.	1016967	16,000.00	168.00	30.24
CHEVROLET	LUV D-MAX 4X2 C/D T/M	2447	18,590.32	909.55	154.62
CHEVROLET	OPTRA 1. 8L T.M.	2163	16,286.22	796.76	135.45
CHEVROLET	LUV D-MAX 4X2 C/D T/M	2797	18,590.32	909.55	154.62
CHEVROLET	LUV D-MAX 4X2 C/D T/M	2708	20,930.00	898.59	152.76
AGENCIA VALLES 2			124,727.00	2,389.41	136.74
NEGOCIOS INDIVIDUALES			27,007.00	334.29	16.23
MAZDA	B2600 C/S STD AC	1013825	12,100.00	108.20	16.23
MAZDA	B2600 C/S STD AC	1017119	14,907.00	226.09	0.00
NEGOCIOS MASIVOS			97,720.00	2,055.12	120.51
TOYOTA	YARIS NITRO	664	12,870.00	123.83	0.00
CHEVROLET	OPTRA 1. 8L T.M.	2118	16,790.00	708.88	120.51
MAZDA	B2600 C/S STD AC	663	18,170.00	339.72	0.00
MAZDA	B2600 C/S STD AC	62	8,500.00	45.82	0.00
NISSAN	FRONTIER AX D/C 4X4	723	27,990.00	691.45	0.00
MAZDA	B2600 C/S STD AC	57	13,400.00	145.42	0.00
TOTAL GENERAL =			5,080,628.69	97,375.79	14,505.48

Figura 3.75 Reporte de Detalle Producción Enero 2005 Última Página

Para verificar los datos en el Data Mart para el mes de Enero 2005, se utilizó la herramienta de generación de información de BusinessObjects Web Intelligence, arrastrando las dimensiones Agencia, Unidad de Negocio y los indicadores Valor Asegurado, Valor Prima Cobrada y Valor Comisión, filtrando únicamente el mes de Enero 2005 como se muestra en la Figura 3.76.

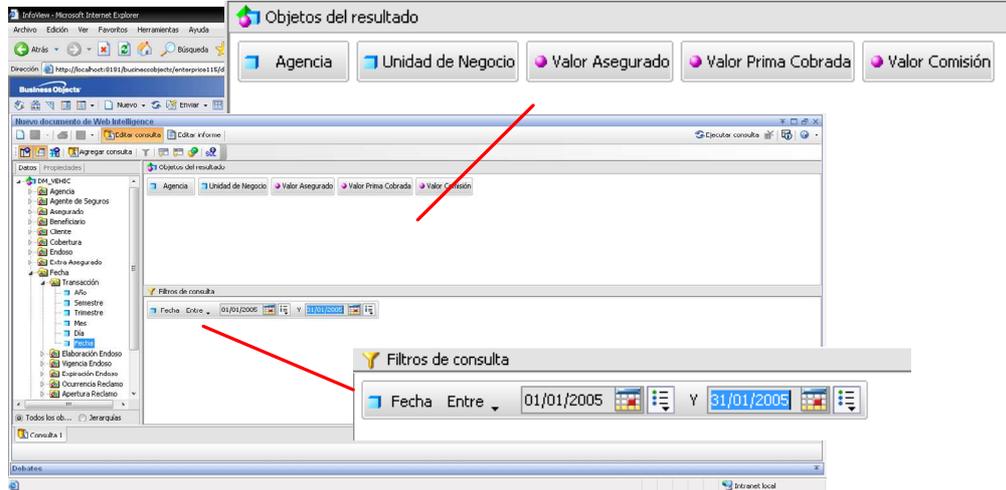


Figura 3.76 Informe de Web Intelligence generado para cuadrar producción de Enero 2005

Al ejecutar la consulta del informe se despliega una tabla organizada de acuerdo al orden de las dimensiones e indicadores arrastrados como se muestra en la Figura 3.77.

Agencia	Unidad de Negocio	Valor Asegurado	Valor Prima Cobrada	Valor Comisión
AGENCIA CENTRO 1	NEGOCIOS INDIVIDUALES	415.693,51	11.824,78	1.650,86
AGENCIA CENTRO 1	NEGOCIOS MASIVOS	61.178,69	632,19	91,59
AGENCIA CENTRO 2	NEGOCIOS INDIVIDUALES	34.931,07	456,16	68,42
AGENCIA CENTRO 2	NEGOCIOS MASIVOS	339.894,49	8.297,94	1.264,45
AGENCIA NORTE 1	NEGOCIOS INDIVIDUALES	617.608,79	7.515,21	893,24
AGENCIA NORTE 1	NEGOCIOS MASIVOS	2.938.722,33	53.085,63	8.443,22
AGENCIA NORTE 2	NEGOCIOS INDIVIDUALES	43.590	1.012,09	151,81
AGENCIA NORTE 2	NEGOCIOS MASIVOS	20.880	393,44	64,9
AGENCIA SUR 1	NEGOCIOS INDIVIDUALES	1.764,1	6	1,02
AGENCIA SUR 1	NEGOCIOS MASIVOS	301.731,85	6.782,22	1.013,03
AGENCIA SUR 2	NEGOCIOS INDIVIDUALES	89.510	1.298,41	99,52
AGENCIA VALLES 1	NEGOCIOS INDIVIDUALES	0	0	0
AGENCIA VALLES 1	NEGOCIOS MASIVOS	90.396,86	3.682,45	627,69
AGENCIA VALLES 2	NEGOCIOS INDIVIDUALES	27.007	334,29	16,23
AGENCIA VALLES 2	NEGOCIOS MASIVOS	97.720	2.055,12	120,51
	Suma:	5.080.628,69	97.375,93	14.506,49

Figura 3.77 Datos de producción generados para Enero 2005 a través de Web Intelligence

Sí se verifica manualmente los valores de valor asegurado, prima cobrada y comisiones se puede observar que coincide con los valores del reporte del

sistema OLTP generado para el mes de Enero 2005, como se muestra en la Figura 3.78.

DETALLE DE PRODUCCIÓN					
RAMO VEHICULOS DESDE 01/01/2005 HASTA 01/31/2005					
<u>MARCA</u>	<u>MODELO</u>	<u>NRO. POLIZA</u>	<u>VALOR ASEGURADO</u>	<u>PRIMA NETA</u>	<u>COMISIÓN</u>
AGENCIA CENTRO 1			476,872.20	12,456.93	1,742.39
NEGOCIOS INDIVIDUALES			415,693.51	11,824.75	1,650.82
HYUNDAI	ACCENT	1007622	6,480.00	165.88	24.88
NEGOCIOS MASIVOS			61,178.69	632.18	91.57
CHEVROLET	CORSA EVOLUCION 1.8S	756	481.60	13.18	1.85
AGENCIA CENTRO 2			374,825.56	8,754.08	1,332.80
NEGOCIOS INDIVIDUALES			34,931.07	456.15	68.42
VOLKSWAGEN	GOL SAVEIRO D.H. 1.8	4	11,190.00	25.22	3.78
NEGOCIOS MASIVOS			339,894.49	8,297.93	1,264.38
CHEVROLET	SPARK 5P STD 1.0 ACTIVI	2274	8,290.00	349.79	59.45
NISSAN	FRONTIER AX D/C 4X4	723	27,990.00	691.45	0.00
MAZDA	B2600 C/S STD AC	57	13,400.00	145.42	0.00
TOTAL GENERAL =			5,080,628.69	97,375.79	14,505.48

Figura 3.78 Reporte de producción de Enero 2005 cortado para ciertas agencias, unidades de negocios y mostrando el gran total.

Como se puede observar los datos coinciden para cada agencia y unidad de negocio con los datos generados a través del Data Mart. Este es un método de cuadro y verificación de datos útil, pero podría ser demoroso. Otro método es consultar directamente sobre las vistas o procedimientos a través de los cuáles son generados los reportes de los sistemas OLTP y unirlos con los datos generados por el Data Mart a través de SQL. Éste método resulta más eficaz para cuadrar gran cantidad de datos, como se muestra a continuación usando una consulta SQL parecida a la siguiente:

```
SELECT  agencia, unidad_negocio, SUM (prima_oltp) AS prima_oltp,
        SUM (prima_datamart) AS prima_datamart,
        SUM (prima_oltp - prima_datamart) AS diff
FROM    (SELECT  agencia, unidad_negocio,
                SUM (produccion_positiva + produccion_negativa
                    ) AS prima_oltp,
                0 AS prima_datamart
        FROM    detalle_produccion
        WHERE   fecha BETWEEN TO_DATE ('01/01/2005', 'DD/MM/YYYY')
                AND TO_DATE ('31/01/2005', 'DD/MM/YYYY')
        GROUP BY agencia, unidad_negocio
        UNION
        SELECT  ag.agencia, un.unidad_negocio, 0 AS prima_oltp,
                SUM (prod.valor_prima_cobrada) AS prima_datamart
        FROM    dim_data.fact_produccion prod,
                dim_data.dim_fecha fec,
                dim_data.dim_agencia ag,
                dim_data.dim_unidad_negocio un
        WHERE   prod.sk_fecha = fec.sk_fecha
```

```

AND prod.sk_agencia = ag.sk_agencia
AND prod.sk_unidad_negocio = un.sk_unidad_negocio
AND fec.anio = 2005
AND fec.mes = 1
GROUP BY ag.agencia, un.unidad_negocio)
GROUP BY agencia, unidad_negocio
ORDER BY agencia, unidad_negocio

```

El SQL utiliza dos tablas virtuales la primera consulta la vista de producción del sistema OLTP y la segunda consulta la tabla de hechos del Data Mart y las dimensiones necesarias. Después se unen y agrupan las dos tablas virtuales generando una columna para los datos agregados del sistema OLTP, otra para los datos agregados del Data Mart y una última para la diferencia.

AGENCIA	UNIDAD_NEGOCIO	PRIMA_OLTP	PRIMA_DATAMART	DIFF
AGENCIA CENTRO 1	NEGOCIOS INDIVIDUALES	11824.75	11824.78	-0.03
AGENCIA CENTRO 1	NEGOCIOS MASIVOS	632.18	632.19	-0.01
AGENCIA CENTRO 2	NEGOCIOS INDIVIDUALES	456.15	456.16	-0.01
AGENCIA CENTRO 2	NEGOCIOS MASIVOS	8297.93	8297.94	-0.01
AGENCIA NORTE 1	NEGOCIOS INDIVIDUALES	7515.17	7515.21	-0.04
AGENCIA NORTE 1	NEGOCIOS MASIVOS	53085.61	53085.63	-0.02
AGENCIA NORTE 2	NEGOCIOS INDIVIDUALES	1012.08	1012.09	-0.01
AGENCIA NORTE 2	NEGOCIOS MASIVOS	393.44	393.44	0
AGENCIA SUR 1	NEGOCIOS INDIVIDUALES	6	6	0
AGENCIA SUR 1	NEGOCIOS MASIVOS	6782.21	6782.22	-0.01
AGENCIA SUR 2	NEGOCIOS INDIVIDUALES	1298.41	1298.41	0
AGENCIA VALLES 1	NEGOCIOS INDIVIDUALES	0	0	0
AGENCIA VALLES 1	NEGOCIOS MASIVOS	3682.45	3682.45	0
AGENCIA VALLES 2	NEGOCIOS INDIVIDUALES	334.29	334.29	0
AGENCIA VALLES 2	NEGOCIOS MASIVOS	2055.12	2055.12	0

Tabla 3.21 Datos obtenidos con el script SQL para cuadro de datos entre el OLTP y el Data Mart

Como se puede observar en la Tabla 3.21 obtenida a través del script SQL de cuadro se obtiene el valor de primas en el OLTP, el valor de primas en el Data Mart y la diferencia entre los dos para cada agencia y unidad de negocio. Como se puede apreciar las diferencias son mínimas en centavos, las mismas que son ocasionadas por precisión de datos. Éste método resulta muy útil para cuadrar datos y realizar las pruebas de verificación de carga de datos necesarias y asegurarse que lo que muestra el sistema OLTP sea consistente con la

información que se cargó en el Data Mart. El mismo método de cuadro usando el script SQL se lo puede modificar para obtener las dimensiones e indicadores necesarios que se deseen cuadrar tanto en el sistema OLTP como en el Data Mart.

3.4.5.2 Cuadre de datos de siniestros

Para realizar el cuadro de datos de siniestros se usó el reporte Detalle Siniestros del sistema OLTP, de igual forma se lo generó para ENERO 2006. El reporte muestra el número de reclamo, la causa de siniestro, el tipo de pérdida y el valor de pagos, gastos y reservas agrupado por agencia y unidad de negocio como se muestra en la Figura 3.79 y Figura 3.80.

DETALLE DE SINIESTROS					
RAMO VEHICULOS DESDE 01/01/2006 HASTA 01/31/2006					
<u>NRO. RECLAMO</u>	<u>CAUSA SINIESTRO</u>	<u>PERDIDA</u>	<u>PAGOS</u>	<u>GASTOS</u>	<u>RESERVAS</u>
AGENCIA CENTRO 1			2,819.50	0.00	3,405.50
NEGOCIOS INDIVIDUALES			2,194.50	0.00	2,705.50
7000015	CHOQUE PARCIAL	PARCIAL	0.00	0.00	-100.00
7000015	CHOQUE PARCIAL	PARCIAL	0.00	0.00	-100.00
7000015	CHOQUE PARCIAL	PARCIAL	0.00	0.00	100.00
7008742	CHOQUE PARCIAL	PARCIAL	0.00	0.00	-4,900.00
7000015	CHOQUE PARCIAL	PARCIAL	0.00	0.00	100.00
7000015	CHOQUE PARCIAL	PARCIAL	0.00	0.00	-100.00
7000015	CHOQUE PARCIAL	PARCIAL	0.00	0.00	100.00
7000015	CHOQUE PARCIAL	PARCIAL	0.00	0.00	100.00
7008742	CHOQUE PARCIAL	PARCIAL	0.00	0.00	-2,194.50
NEGOCIOS MASIVOS			625.00	0.00	700.00
7003440	CHOQUE PARCIAL	PARCIAL	0.00	0.00	-700.00
7003440	CHOQUE PARCIAL	PARCIAL	0.00	0.00	700.00

Figura 3.79 Reporte de Detalle Siniestros Enero 2006 Primera Página

Código	Descripción	Estado	Valor Pago	Valor Gasto	Valor Reservas Constituidas	Valor Reservas Liberadas
7002980	CHOQUE PARCIAL	PARCIAL	0.00	0.00	-127.11	
7000949	CHOQUE PARCIAL	PARCIAL	0.00	0.00	100.00	
NEGOCIOS MASIVOS			812.42	0.00	-2,258.18	
7000956	CHOQUE PARCIAL	PARCIAL	0.00	0.00	100.00	
7008834	CHOQUE PARCIAL	PARCIAL	0.00	0.00	-355.51	
7009515	CHOQUE PARCIAL	PARCIAL	0.00	0.00	0.00	
7009515	CHOQUE PARCIAL	PARCIAL	0.00	0.00	-1,003.43	
7008834	CHOQUE PARCIAL	PARCIAL	0.00	0.00	0.00	
7000956	CHOQUE PARCIAL	PARCIAL	0.00	0.00	100.00	
7007914	CHOQUE PARCIAL	PARCIAL	0.00	0.00	1,900.00	
7009515	CHOQUE PARCIAL	PARCIAL	0.00	0.00	-495.00	
7008834	CHOQUE PARCIAL	PARCIAL	0.00	0.00	-504.24	
7007914	CHOQUE PARCIAL	PARCIAL	0.00	0.00	-2,000.00	
7007914	CHOQUE PARCIAL	PARCIAL	0.00	0.00	0.00	
TOTAL GENERAL =			25,373.59	1,418.50	34,367.75	

Figura 3.80 Reporte de Detalle Siniestros Enero 2006 Última Página

El cuadro de datos para el mes de Enero 2006 correspondiente a siniestros se lo realizó de la misma forma que el cuadro de datos para producción. Se generó un informe usando Web Intelligence para el mes de Enero 2006 arrastrando las dimensiones, agencia, unidad de negocio y los indicadores valor pago, valor gasto, valor reservas constituidas y valor reservas liberadas como se muestra en la Figura 3.81.

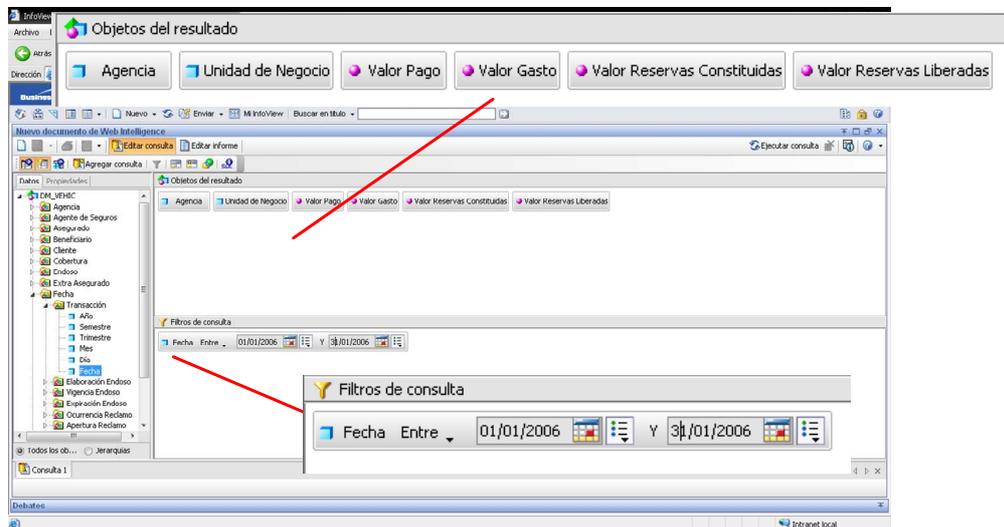


Figura 3.81 Informe de Web Intelligence para cuadro de siniestros

Al ejecutar la consulta del informe de Web Intelligence se muestra de forma tabular los datos arrastrados. Se cambio la forma de presentación de los datos para obtener la agencia, unidad de negocio, valor de pagos, gastos y la suma de reservas constituidas y liberadas y poderlos comprobar con el reporte de Detalle Siniestros para Enero 2006 del sistema OLTP como se muestra en la Tabla 3.22.

Agencia	Unidad de Negocio	Valor Pago	Valor Gasto	Reservas
AGENCIA CENTRO 1	NEGOCIOS INDIVIDUALES	2.194,5	0	10.400
AGENCIA CENTRO 1	NEGOCIOS MASIVOS	625	0	2.800
AGENCIA CENTRO 2	NEGOCIOS MASIVOS	10.607,49	0	87.710
AGENCIA NORTE 1	NEGOCIOS INDIVIDUALES	2.990	0	52.692
AGENCIA NORTE 1	NEGOCIOS MASIVOS	6.341,17	1.418,5	269.228,22
AGENCIA NORTE 2	NEGOCIOS INDIVIDUALES	0	0	1.150,02
AGENCIA NORTE 2	NEGOCIOS MASIVOS	0	0	800
AGENCIA SUR 1	NEGOCIOS MASIVOS	50,4	0	6.040
AGENCIA SUR 2	NEGOCIOS INDIVIDUALES	0	0	200
AGENCIA SUR 2	NEGOCIOS MASIVOS	0	0	3.200
AGENCIA VALLES 1	NEGOCIOS MASIVOS	1.752,61	0	2.700
AGENCIA VALLES 2	NEGOCIOS INDIVIDUALES	0	0	202
AGENCIA VALLES 2	NEGOCIOS MASIVOS	812,42	0	4.200
	Suma:	25.373,59	1.418,5	441.322,24

Tabla 3.22 Datos de siniestros para Enero 2006 generados a través de Web Intelligence

El método de cuadro de siniestros fue similar al utilizado para cuadrar producción, verificando manualmente que los valores para cada agencia y unidad de negocio coincidan con los datos obtenidos a través del Data Mart. Como se mencionó anteriormente éste método resultado muy demandante con respecto al tiempo y puede ser propenso a errores, se recomienda utilizar el método del script SQL de cuadro consultando las vistas y procedimientos del sistema y uniendo y agrupando con los datos cargados en el Data Mart tal como se explica en el Capítulo 3.4.5.1.

CAPÍTULO IV

4 APLICACIÓN DEL DATA MART AL CASO DE ESTUDIO

Una vez construido el Data Mart para el análisis del ramo de vehículos para una empresa aseguradora es necesario mostrar la información generada por el mismo para que sea analizada a través de reportes, tableros o minería de datos.

El alcance de éste proyecto de titulación solo abarca la presentación de información del Data Mart para análisis a través de reportes claves del negocio de seguros y un tablero de control.

La plataforma de inteligencia de negocios BusinessObjects posee el portal Web de entrega de información InfoView. A través de InfoView se accederá a toda la información generada por el Data Mart. Además, InfoView constituye la interfaz para acceder a la herramienta Web Intelligence que permitirá a los usuarios del negocio generar información a demanda y la herramienta Performance Management que se encargará de gestionar las métricas, analíticas y el tablero de control construido.

4.1 DEFINICIÓN DE LOS INDICADORES DE GESTIÓN

En la Tabla 4.1 se describen los indicadores de gestión de seguros más relevantes para el análisis del ramo de vehículos.

Indicador	Fórmula	Definición
Prima Cobrada	=[Valor Prima Cobrada]	<p>Valor monetario que se ha cobrado a los clientes de las pólizas.</p> <p>Representa un ingreso para la Compañía y su signo es positivo.</p>

Pagos	$= [\text{Valor Pago}] * -1$	Valor monetario que se ha pagado al beneficiario del seguro por concepto de un siniestro cubierto. Representa un egreso para la Compañía y su signo es negativo.
Gastos	$= [\text{Valor Gasto}] * -1$	Valor monetario que representa los gastos que ha ocasionado un siniestro como por ejemplo el pago a un verificador del siniestro. Representa un egreso para la Compañía y su signo es negativo.
Siniestros Pagados	$= ([\text{Valor Pago}] + [\text{Valor Gasto}]) * -1$	Valor monetario que representa el pago total de un siniestro incluyendo los gastos que esto ocasionó.
Reservas Constituidas	$= [\text{Valor Reservas Constituidas}]$	Valor monetario que se ha constituido para hacer frente a un siniestro reportado por un cliente. Representa un valor que la Compañía se reserva para hacer frente al pago total de un siniestro. La constitución es un valor positivo.
Reservas Liberadas	$= [\text{Valor Reservas Liberadas}]$	Valor monetario que se libera cuando se ha concluido el pago de un

		siniestro. Representa un valor que la Compañía libera cuando se haya o no cerrado un reclamo que ocasionó un siniestro. La liberación es un valor negativo.
Reservas	= [Valor Reservas Constituidas] + [Valor Reservas Liberadas]	Valor monetario neto resultante de las reservas constituidas y liberadas.
Comisiones Pagadas	= [Valor Comisión] * -1	Valor monetario que se ha pagado por concepto de comisiones a un Agente de Seguros. Representa un egreso para la Compañía y su signo es negativo.
Resultado Técnico	= [Valor Prima Cobrada] + [Siniestros Pagados] + [Reservas] + [Comisiones Pagadas]	Valor monetario que representa los resultados finales de la Compañía. Incluye los ingresos que representan las primas cobradas y los egresos representados por los siniestros pagados, reservas y comisiones pagadas. Es un indicador propio del negocio asegurador.
Valor Asegurado	= [Valor Asegurado]	Valor monetario que ha sido asegurado en toda la cartera de pólizas cobradas.
Número de vehículos	= Cuenta([Motor];	Representa el número de

asegurados.	ValoresÚnicos)	vehículos que han sido asegurados por la Compañía.
Valor Asegurado Promedio	= [Valor Asegurado] / [Nro. Vehículos Asegurados]	Valor monetario que representa el valor asegurado promedio para toda la cartera de vehículos asegurados.
Tasa	= [Valor Prima Cobrada] / [Valor Asegurado]	Representa la relación entre el valor monetario de primas cobradas y el valor asegurado. En base a la tasa se calcula la prima que será cobrada para un vehículo asegurado. Por ejemplo, si la tasa es del 4% y el valor asegurado del vehículo es de \$10000 el valor de primas a cobrar será de \$400.
% Siniestralidad	= [Siniestros Pagados] / [Valor Prima Cobrada]	Representa el porcentaje que ha sido pagado en siniestros del valor total de primas cobradas. La siniestralidad es un indicador clave para evaluar como se están suscribiendo nuevas pólizas de vehículos.
% Comisiones Pagadas	= [Valor Comisión] / [Valor Prima Cobrada]	Representa el porcentaje de las primas cobradas que se están pagando por concepto de

		comisiones a los agentes de seguros.
Número de vehículos siniestrados	=Cuenta([Nro. Reclamo]; ValoresÚnicos)	Número de vehículos que se han siniestrado.
Frecuencia	=[Nro. Vehículos Siniestrados] / [Nro. Vehículos Asegurados]	Relación entre el número de vehículos siniestrados y el número de vehículos asegurados. Representa el porcentaje de vehículos asegurados que se ha siniestrado.

Tabla 4.1 Indicadores de Gestión de Seguros

4.2 REPORTES CON LOS INDICADORES DE GESTIÓN

Se crearon cuatro reportes que responden a todas las preguntas del negocio definidas en los requerimientos funcionales. Los reportes fueron construidos usando la herramienta Web Intelligence y son accedidos a través del portal Web de entrega de información InfoView como se muestra en la Figura 4.1.

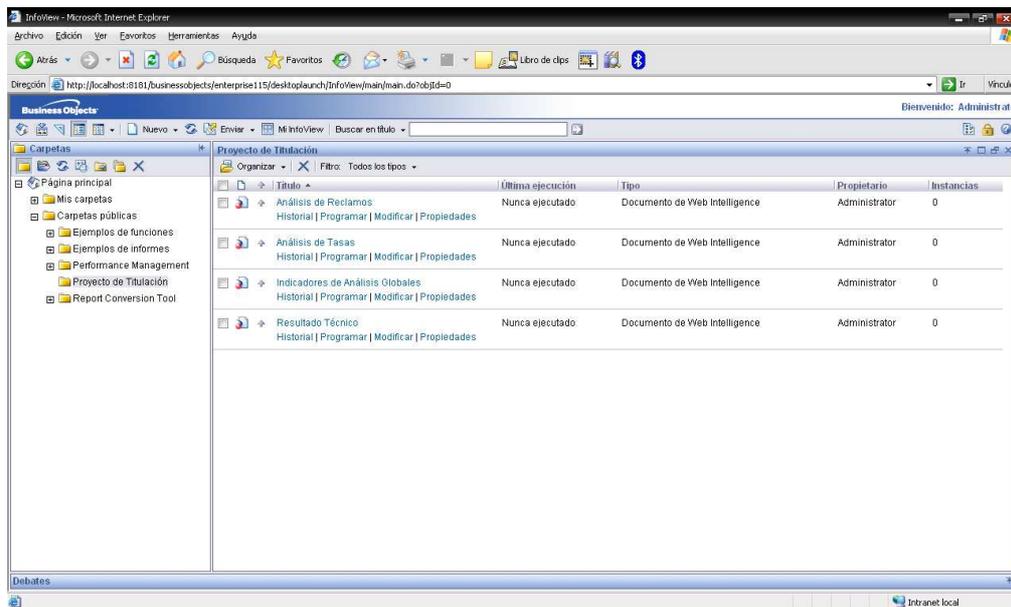


Figura 4.1 Portal de entrega de información BusinessObjects InfoView

Web Intelligence e InfoView tienen la opción de Exploración que permite delimitar el contexto de análisis en un reporte. Los reportes elaborados para el proyecto de titulación muestran inicialmente todos los datos pero a través de la opción de exploración se puede analizar por año, agencia, agente de seguros, marca, modelo, etc. respondiéndose a las preguntas del negocio. Se elaboraron los siguientes reportes de análisis descritos en la Tabla 4.2.

Reporte	Descripción
Resultado Técnico	Muestra los ingresos, egresos y el resultado técnico de la Compañía para que pueda ser analizado por año, mes, agencia, unidad de negocio, agente de seguros y cliente. Además, muestra los ingresos por concepto de primas y los egresos por concepto de siniestros.
Indicadores de Análisis Globales	En éste reporte se consolidan los indicadores de gestión de seguros de vehículos más relevantes. Se incluyen todos los indicadores que ayudan a monitorear y evaluar la salud de un ramo por año, mes, agencia, unidad de negocio, agente de seguros, cliente, marca de vehículo, modelo, etc.
Análisis de Tasas	El reporte de análisis de tasas permite evaluar las tasas que se han cobrado por mes y año. Además, provee un resumen del número de vehículos asegurados, valor asegurado, valor asegurado promedio, prima cobrada, prima promedio y tasa por marca. En éste reporte se elaboró una jerarquía que permite hacer un drill-down desde marca hacia modelo.
Análisis de Reclamos	El reporte de análisis de reclamos permite analizar el número de vehículos siniestrados y el valor reclamo promedio por agencia, año y su tendencia. También permite analizar la tendencia en el número de vehículos siniestrados por año, mes y causa de forma tabular y gráfica.

Tabla 4.2 Reportes de análisis construidos

4.2.1 REPORTE DE RESULTADO TÉCNICO

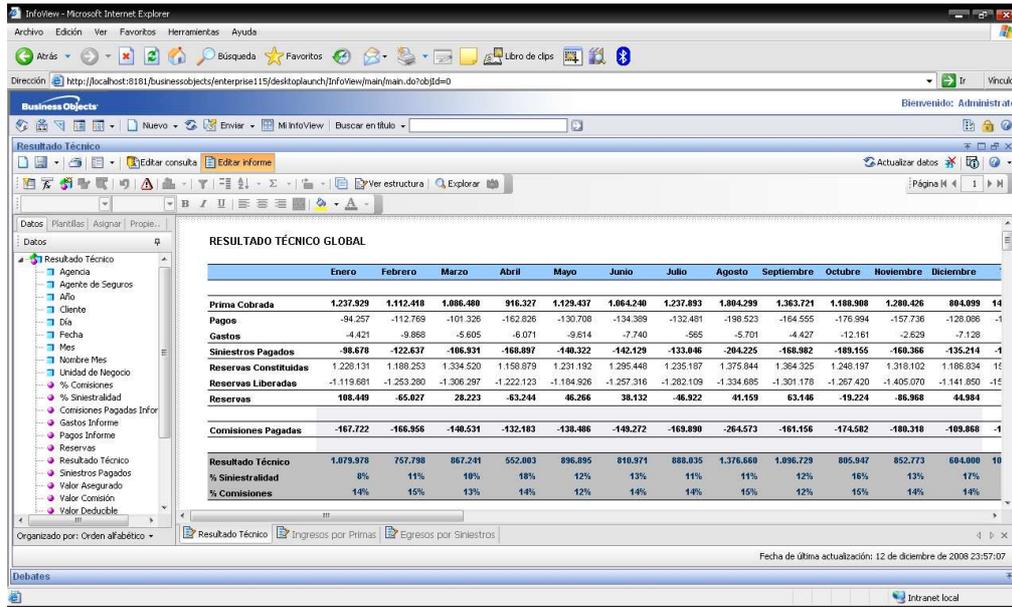


Figura 4.2 Pantalla de acceso al reporte de Resultado Técnico

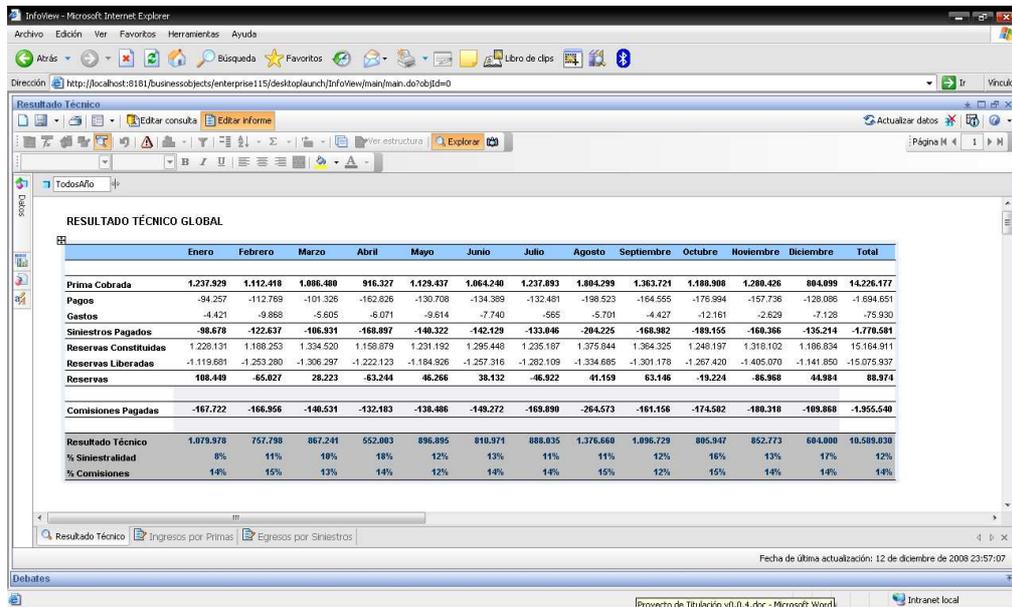


Figura 4.3 Resultado Técnico para todos los años (2005 – 2007)

	Enero	Febrero	Marzo	Abril	Mayo	Junio	Julio	Agosto	Septiembre	Octubre	Noviembre	Diciembre	Total
Prima Cobrada	97.376	85.761	84.395	83.785	168.788	166.444	198.953	238.669	192.463	183.903	195.784	131.377	1.821.788
Pagos	-9.158	-940	-15.746	-8.859	-10.955	-5.820	-30.460	-33.926	-14.077	-17.449	-4.966	-4.405	-156.793
Gastos	0	-600	0	0	-235	-330	0	-130	-50	-4.377	-275	-475	-6.472
Sinistros Pagados	-9.158	-1.540	-15.746	-8.859	-11.230	-6.150	-30.460	-34.056	-14.127	-21.826	-5.231	-4.880	-163.265
Reservas Constituidas	94.001	79.969	103.711	84.788	97.821	118.032	153.709	198.004	136.251	182.749	187.669	229.787	1.666.479
Reservas Liberadas	-75.474	-81.847	-99.915	-84.417	-99.759	-98.314	-177.197	-174.811	-149.581	-193.268	-153.939	-205.407	-1.593.928
Reservas	18.526	-1.879	3.796	371	-1.938	19.717	-23.488	23.193	-13.330	-10.519	33.720	24.379	72.551
Comisiones Pagadas	-14.586	-12.432	-13.868	-12.327	-22.210	-25.806	-33.262	-38.614	-30.400	-29.638	-28.416	-18.746	-278.624
Resultado Técnico	92.238	69.911	69.377	62.971	125.420	165.886	111.743	181.191	134.687	121.920	195.856	132.129	1.462.369
% Sinistralidad	9%	2%	1%	11%	7%	4%	15%	15%	7%	12%	3%	4%	9%
% Comisiones	15%	14%	14%	15%	14%	15%	17%	17%	16%	16%	15%	14%	15%

Figura 4.4 Resultado Técnico explorando el año 2005

	Enero	Febrero	Marzo	Abril	Mayo	Junio	Julio	Agosto	Septiembre	Octubre	Noviembre	Diciembre	Total
Prima Cobrada	394.674	372.888	366.797	382.655	485.888	389.468	464.582	532.489	558.275	412.838	591.236	342.868	5.294.878
Pagos	-25.374	-25.472	-12.720	-40.839	-49.655	-34.592	-23.980	-55.667	-46.983	-40.167	-58.630	-59.250	-473.338
Gastos	-1.419	-4.538	-270	-1.674	-4.164	-508	-480	-150	-2.213	-4.834	-679	-3.921	-24.870
Sinistros Pagados	-26.792	-30.889	-12.988	-42.512	-53.829	-35.121	-24.460	-55.817	-49.196	-45.081	-59.389	-43.171	-498.288
Reservas Constituidas	220.661	374.070	294.718	386.307	421.427	435.889	415.743	502.339	496.675	437.366	576.582	518.908	5.050.686
Reservas Liberadas	-186.293	-401.911	-290.482	-420.513	-400.831	-418.824	-422.223	-494.657	-493.065	-444.715	-618.888	-507.826	-5.100.309
Reservas	34.368	-27.841	4.236	-54.206	28.596	16.946	-6.480	7.682	-6.389	-7.350	-42.296	11.882	-49.643
Comisiones Pagadas	-64.091	-56.918	-58.656	-55.284	-61.283	-63.296	-71.369	-73.383	-64.056	-64.735	-79.377	-47.820	-758.229
Resultado Técnico	338.246	258.219	299.386	238.733	318.572	387.996	382.282	411.851	438.633	295.744	418.264	243.751	3.998.798
% Sinistralidad	7%	8%	4%	11%	13%	9%	5%	10%	9%	11%	10%	10%	10%
% Comisiones	16%	15%	16%	14%	15%	16%	15%	14%	12%	16%	13%	14%	15%

Figura 4.5 Resultado Técnico explorando el año 2006

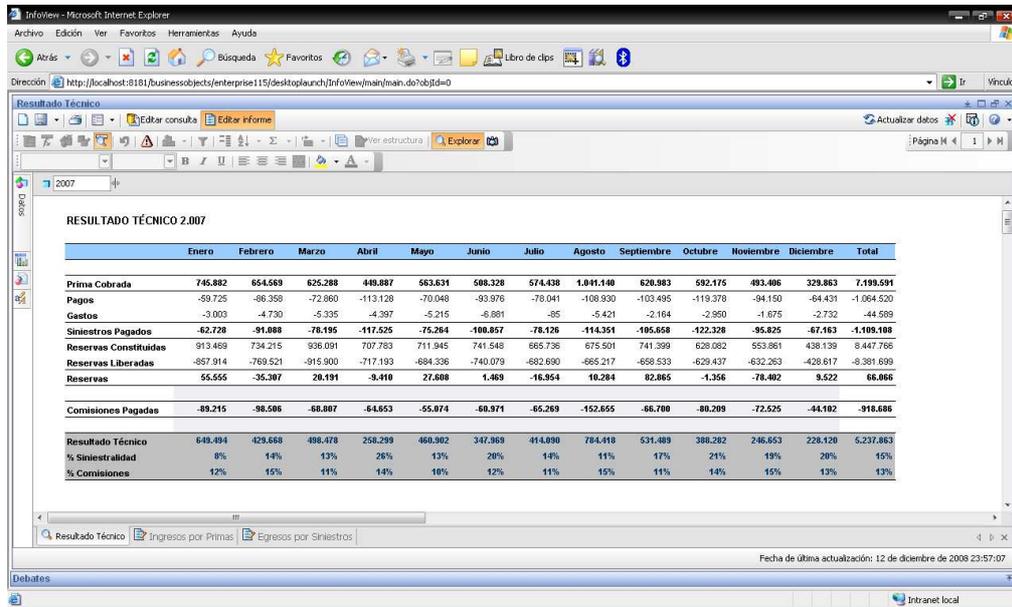


Figura 4.6 Resultado Técnico explorando el año 2007

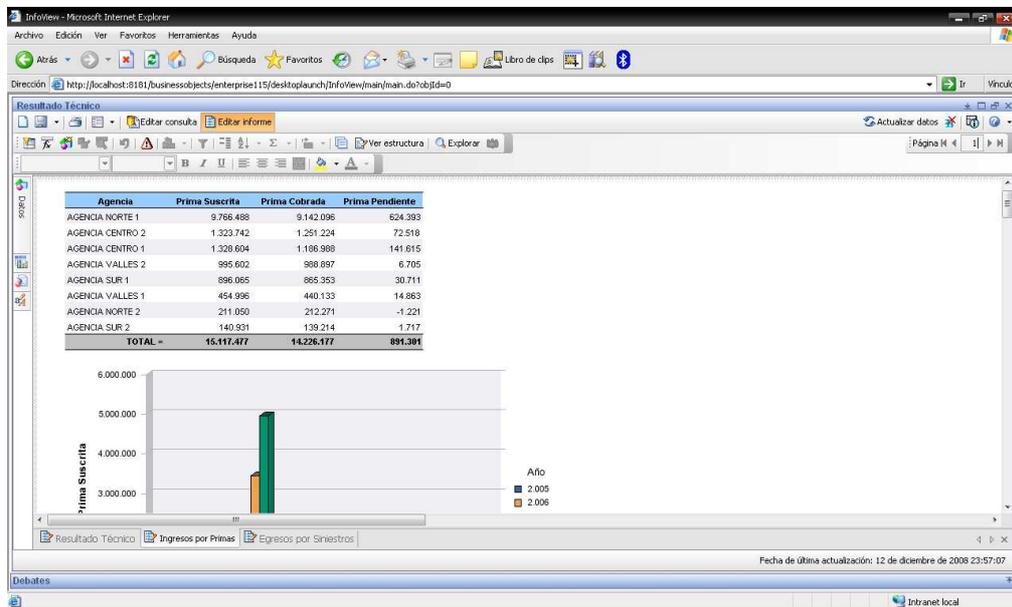


Figura 4.7 Informe de ingresos por primas para todos los años

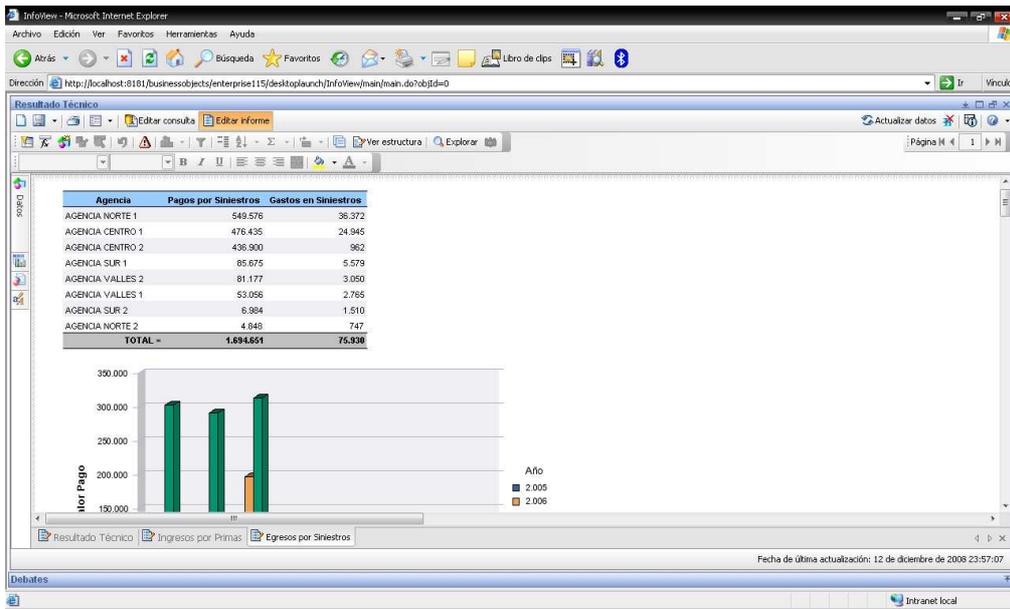


Figura 4.8 Informe de egresos por siniestros para todos los años

4.2.2 REPORTE DE INDICADORES DE ANÁLISIS GLOBALES

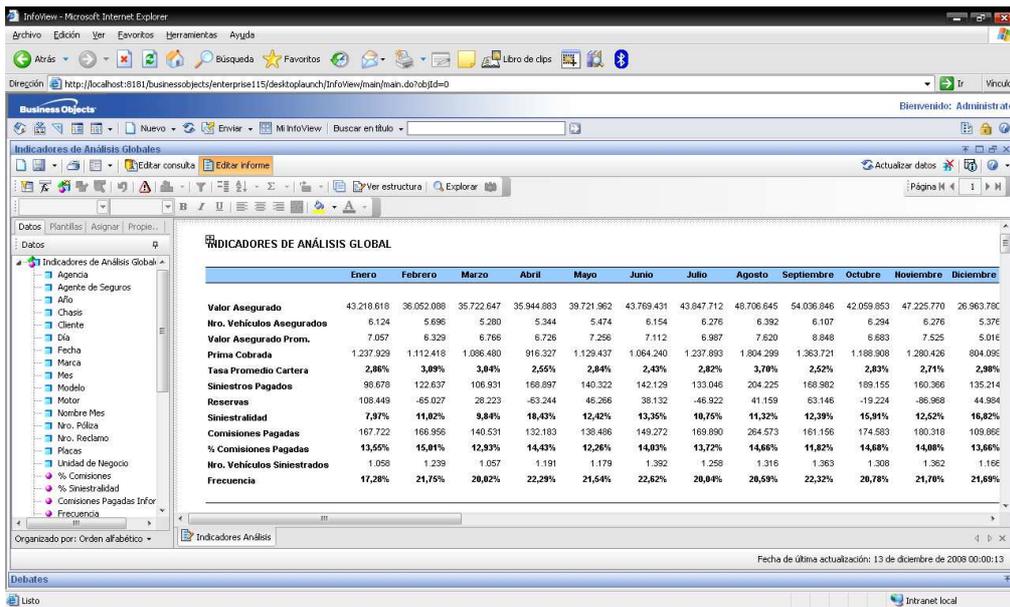


Figura 4.9 Pantalla de acceso al reporte de indicadores de análisis globales

Indicadores de Análisis Globales

	Enero	Febrero	Marzo	Abril	Mayo	Junio	Julio	Agosto	Septiembre	Octubre	Noviembre	Diciembre	Total
Valor Asegurado	43.218.618	36.052.088	35.722.647	35.944.883	39.721.962	43.769.431	43.847.712	48.706.645	54.036.846	42.059.853	47.225.770	26.963.780	497.270.234
Nro. Vehículos Asegurados	6.124	5.696	5.280	5.344	5.474	6.154	6.276	6.392	6.107	6.294	6.276	5.376	20.769
Valor Asegurado Prom.	7.057	6.329	6.766	6.726	7.256	7.112	6.987	7.620	8.848	6.683	7.525	5.016	23.943
Prima Cobrada	1.237.929	1.112.418	1.066.480	916.327	1.129.437	1.064.240	1.237.893	1.804.289	1.363.721	1.188.908	1.280.426	604.099	14.226.177
Tasa Promedio Cartera	2,86%	3,89%	3,04%	2,55%	2,84%	2,43%	2,82%	3,70%	2,52%	2,83%	2,71%	2,98%	2,86%
Sinistros Pagados	98.678	122.637	106.931	168.897	140.322	142.129	133.046	204.225	168.982	189.155	160.366	135.214	1.770.581
Reservas	108.449	-85.027	28.223	-83.244	46.266	38.132	-46.922	41.159	63.146	-19.224	-86.968	44.984	88.974
Siniestralidad	7,97%	11,02%	9,84%	18,43%	12,42%	13,35%	16,75%	11,32%	12,39%	15,91%	12,52%	16,82%	12,45%
Comisiones Pagadas	167.722	166.956	140.531	132.183	138.486	149.272	169.890	264.573	161.156	174.583	180.318	109.868	1.955.540
% Comisiones Pagadas	13,55%	15,01%	12,93%	14,43%	12,26%	14,03%	13,72%	14,66%	11,82%	14,68%	14,08%	13,66%	13,75%
Nro. Vehículos Sinistrados	1.058	1.239	1.057	1.191	1.179	1.392	1.258	1.316	1.363	1.308	1.362	1.166	7.505
Frecuencia	17,28%	21,75%	20,02%	22,29%	21,54%	22,62%	20,04%	20,59%	22,32%	20,78%	21,70%	21,69%	36,14%

Fecha de última actualización: 13 de diciembre de 2008 00:00:13

Figura 4.10 Indicadores de análisis globales para todos los años

Indicadores de Análisis 2.006 - AGENCIA ASESORA PRODUCTORA DE SEGUROS 17

	Enero	Febrero	Marzo	Abril	Mayo	Junio	Julio	Agosto	Septiembre	Octubre	Noviembre	Diciembre	Total
Valor Asegurado	1.302.870	1.422.689	1.770.937	2.135.575	2.206.695	1.981.413	2.459.194	3.571.924	5.446.238	3.321.205	4.350.701	4.027.503	33.996.933
Nro. Vehículos Asegurados	261	276	266	291	319	300	576	533	546	413	470	546	2.102
Valor Asegurado Prom.	4.992	5.155	6.658	7.339	6.918	6.605	4.269	6.702	9.975	8.042	9.257	7.376	16.174
Prima Cobrada	19.564	27.032	30.930	33.402	43.902	35.201	46.596	78.753	129.367	67.903	112.712	129.860	755.222
Tasa Promedio Cartera	1,59%	1,99%	1,75%	1,56%	1,89%	1,78%	1,89%	2,20%	2,38%	2,84%	2,59%	3,22%	2,22%
Sinistros Pagados	652	5.259	6.657	536	4.463	3.256	11.406	16.455	19.755	2.234	8.430	13.172	86.274
Reservas	5.866	-4.168	1.015	-2.910	-4.837	-2.871	3.651	423	-324	-3.817	4.979	-1.297	-4.289
Siniestralidad	3,33%	19,45%	21,52%	1,60%	21,55%	8,25%	24,48%	20,89%	14,59%	3,29%	7,40%	10,14%	12,75%
Comisiones Pagadas	3.040	3.885	4.423	4.745	6.254	4.943	6.575	10.522	15.654	9.582	15.845	18.497	103.963
% Comisiones Pagadas	15,64%	14,37%	14,30%	14,20%	14,24%	14,04%	14,11%	13,36%	12,10%	14,11%	14,06%	14,24%	13,77%
Nro. Vehículos Sinistrados	80	105	83	77	74	91	79	86	86	77	78	75	553
Frecuencia	22,99%	38,04%	31,20%	26,46%	23,20%	30,33%	13,72%	14,26%	15,75%	18,64%	16,60%	13,74%	26,31%

Fecha de última actualización: 13 de diciembre de 2008 00:00:13

Figura 4.11 Indicadores de análisis globales para el año 2006 y el agente de seguros AGENCIA ASESORA PRODUCTORA DE SEGUROS 17

INDICADORES DE ANÁLISIS 2.007 - AGENCIA NORTE 1

	Enero	Febrero	Marzo	Abril	Mayo	Junio	Julio	Agosto	Septiembre	Octubre	Noviembre	Diciembre	Total
Valor Asegurado	15,734,897	13,563,363	11,588,243	13,384,223	10,438,220	12,331,343	14,415,066	11,683,794	16,476,347	13,169,358	14,866,174	5,016,508	152,847,538
Nro. Vehículos Asegurados	3,127	2,510	2,495	2,468	2,352	2,839	2,857	2,383	2,811	2,465	2,405	1,511	9,760
Valor Asegurado Prom.	5,032	5,404	4,645	5,415	4,438	4,344	5,046	4,903	6,310	5,343	6,181	3,320	15,640
Prima Cobrada	476,644	406,997	399,910	281,900	356,401	318,528	368,731	894,166	389,752	342,991	327,870	128,721	4,692,610
Tasa Promedio Cartera	3,03%	3,00%	3,45%	2,11%	3,41%	2,58%	2,56%	7,65%	2,37%	2,60%	2,21%	2,57%	3,07%
Sinistros Pagados	26,347	22,606	16,888	32,411	23,390	25,805	19,051	43,990	53,196	29,697	15,490	22,639	331,511
Reservas	9,823	-5,894	11,651	-3,124	2,786	7,904	2,810	7,691	5,040	-572	812	2,315	41,244
Sinistralidad	5,53%	5,55%	4,22%	11,50%	6,56%	8,10%	5,17%	4,92%	13,65%	8,66%	4,72%	17,59%	7,06%
Comisiones Pagadas	53,786	61,395	43,318	39,640	34,968	37,544	43,984	131,489	43,067	51,586	48,148	18,759	607,693
% Comisiones Pagadas	11,28%	15,88%	10,83%	14,66%	9,81%	11,79%	11,93%	14,71%	11,05%	15,04%	14,68%	14,57%	12,95%
Nro. Vehículos Sinistrados	502	361	406	410	367	411	368	396	367	385	443	319	3,013
Frecuencia	16,05%	14,38%	16,27%	16,61%	15,69%	14,48%	12,81%	16,62%	14,06%	15,62%	18,42%	21,11%	30,87%

Fecha de última actualización: 13 de diciembre de 2008 00:00:13

Figura 4.12 Indicadores de análisis globales para el año 2007 y la AGENCIA NORTE 1

INDICADORES DE ANÁLISIS 2.007 - CHEVROLET

	Enero	Febrero	Marzo	Abril	Mayo	Junio	Julio	Agosto	Septiembre	Octubre	Noviembre	Diciembre	Total
Valor Asegurado	11,204,366	11,122,336	9,920,199	9,309,760	6,198,804	7,833,356	6,157,211	7,317,744	6,644,623	6,757,290	7,649,195	2,648,995	91,763,838
Nro. Vehículos Asegurados	1,783	1,478	1,307	1,203	1,236	1,361	1,192	1,108	1,044	1,134	1,119	957	6,732
Valor Asegurado Prom.	6,284	7,525	6,825	7,739	5,015	5,756	5,165	6,604	6,365	5,969	6,836	2,768	13,631
Prima Cobrada	334,163	363,495	271,147	187,705	179,048	166,521	180,196	801,195	185,701	228,909	216,682	75,286	3,150,047
Tasa Promedio Cartera	2,98%	3,27%	3,04%	2,82%	2,89%	2,13%	2,69%	10,95%	2,49%	3,39%	2,83%	2,84%	3,43%
Sinistros Pagados	37,878	57,833	35,222	54,615	51,296	27,206	46,424	72,975	53,182	52,974	40,056	37,192	566,852
Reservas	25,091	-17,229	24,271	-12,870	9,051	4,459	-10,210	3,267	2,318	-4,793	1,535	2,842	27,743
Sinistralidad	11,34%	15,91%	12,89%	29,10%	28,65%	16,34%	28,98%	8,11%	32,10%	23,14%	18,49%	49,40%	18,00%
Comisiones Pagadas	51,775	58,035	40,920	28,645	26,784	24,687	24,588	119,730	25,699	37,717	35,826	12,021	486,426
% Comisiones Pagadas	15,49%	15,97%	15,09%	15,26%	14,96%	14,83%	15,35%	14,94%	15,51%	16,48%	16,53%	15,97%	15,44%
Nro. Vehículos Sinistrados	410	314	328	326	297	295	268	275	259	262	298	227	2,308
Frecuencia	22,99%	21,24%	25,10%	27,10%	24,03%	21,68%	22,48%	24,82%	24,81%	23,10%	26,63%	23,72%	34,28%

Fecha de última actualización: 13 de diciembre de 2008 00:00:13

Figura 4.13 Indicadores de análisis globales para el año 2007 y la marca CHEVROLET

4.2.3 REPORTE DE ANÁLISIS DE TASAS

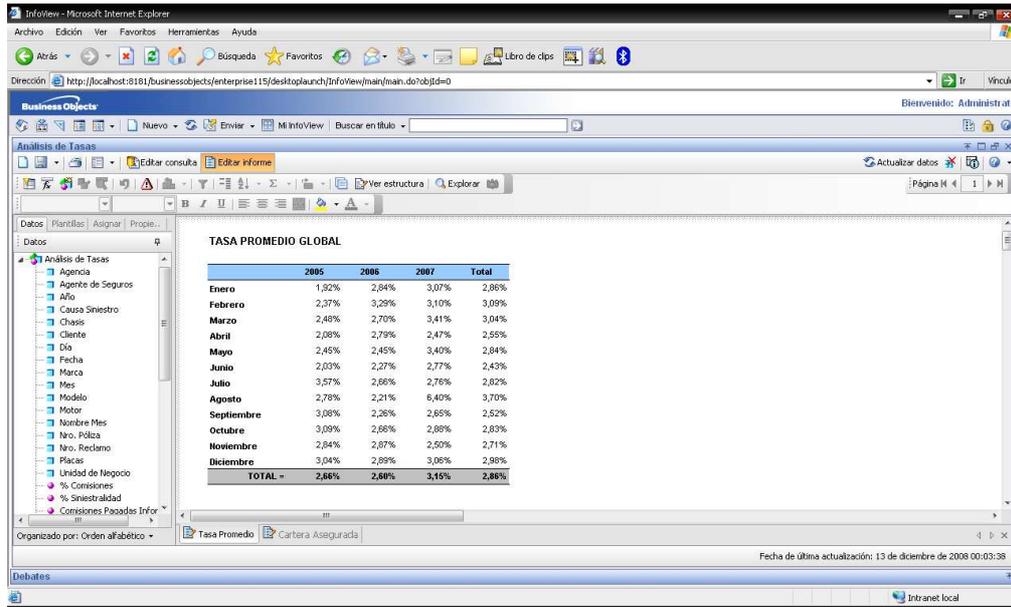


Figura 4.14 Pantalla de acceso al reporte de análisis de tasas

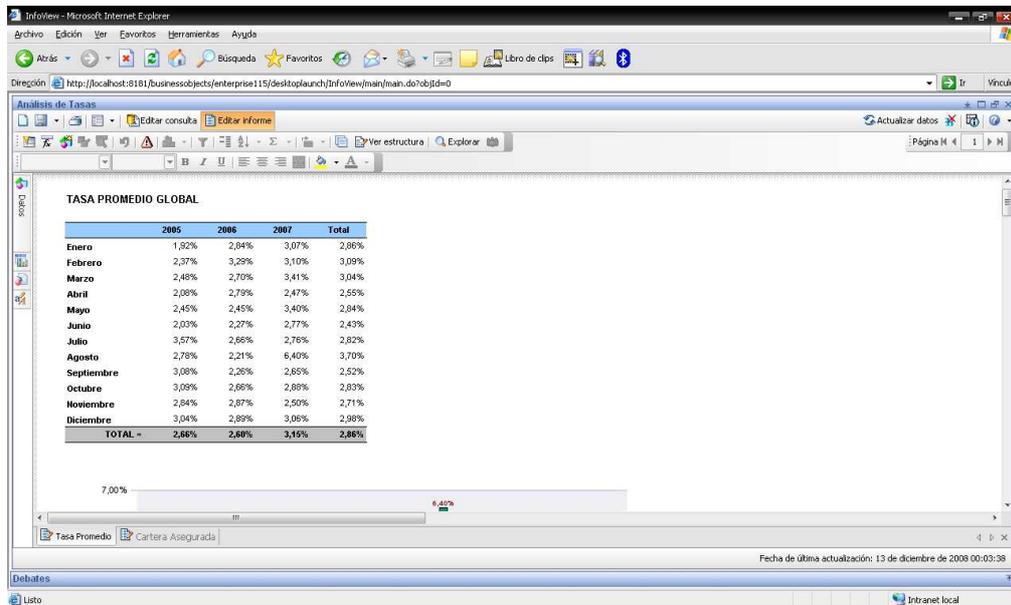


Figura 4.15 Tasa Promedio por año y mes de forma tabular

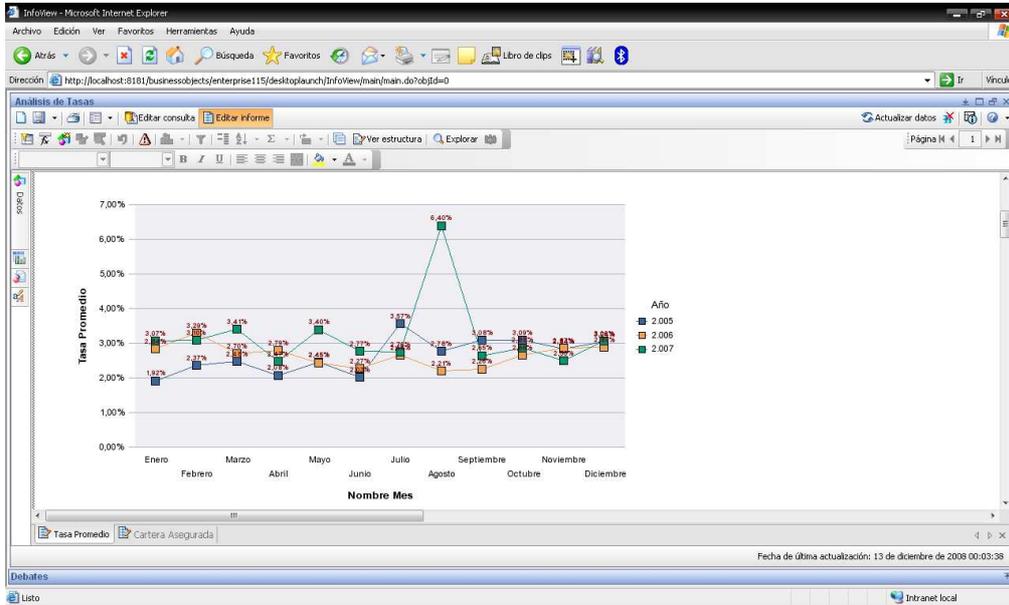


Figura 4.17 Tasa promedio por marca de vehículo

	Nro. Vhs Aseg.	Valor Asegurado	Valor Aseg. Prom.	Prima Cobrada	Prima Promedio	Tasa
CHEVROLET	11053	236.399.239	21.388	7.533.903	682	3,19%
MAZDA	2102	51.221.876	24.368	1.352.004	643	2,64%
HYUNDAI	1500	44.310.733	29.540	1.097.588	732	2,48%
KIA	963	23.836.269	24.752	596.224	619	2,60%
NISSAN	861	21.990.361	33.268	565.235	855	2,57%
VOLKSWAGEN	1070	22.006.417	20.567	454.601	425	2,07%
TOYOTA	597	13.657.914	22.678	339.456	569	2,49%
MITSUBISHI	380	12.279.484	32.314	316.597	833	2,58%
FORD	354	12.180.656	34.409	311.077	879	2,55%
INTERNATIONAL	75	8.829.223	117.723	299.752	3.997	3,40%
PEUGEOT	454	10.482.733	23.112	238.382	525	2,27%
KENWORTH	32	4.076.428	127.388	138.774	4.337	3,40%
MERCEDES BENZ	51	3.607.861	70.742	122.303	2.398	3,39%
VOLVO	26	3.322.353	127.783	117.319	4.512	3,53%
HINO	65	3.582.307	55.112	112.754	1.735	3,15%
CITROEN	223	6.051.929	27.139	109.602	491	1,81%
RENAULT	365	4.526.013	12.400	96.568	265	2,13%

InfoView - Microsoft Internet Explorer

Archivo Edición Ver Favoritos Herramientas Ayuda

Dirección http://localhost:8181/businessobjects/enterprise115/desktoplaunch/InfoView/main/main.do?objId=0

Analisis de Tasas

2007

TASA POR MARCA 2.007

	Nro. Vhs Aseg.	Valor Asegurado	Valor Aseg. Prom.	Prima Cobrada	Prima Promedio	Tasa
CHEVROLET	6732	91.763.838	13.631	3.150.047	468	3,43%
MAZDA	1815	27.520.736	15.163	879.465	485	3,20%
HYUNDAI	1284	20.962.744	16.326	636.041	495	3,03%
KIA	881	12.309.087	13.972	358.266	407	2,91%
NISSAN	559	10.274.429	18.380	316.220	566	3,08%
TOYOTA	572	9.917.232	17.338	272.169	476	2,74%
FORD	300	8.459.631	28.199	225.198	751	2,66%
MITSUBISHI	348	7.891.579	22.677	220.621	634	2,80%
VOLKSWAGEN	738	9.091.175	10.964	205.685	279	2,54%
INTERNATIONAL	75	5.633.757	75.117	188.046	2.507	3,34%
PEUGEOT	394	5.177.722	13.141	138.467	346	2,64%
KENWORTH	29	2.843.755	91.164	96.008	3.311	3,63%
RENAULT	354	3.563.423	10.066	92.137	232	2,31%
HINO	44	1.633.257	37.119	67.903	1.316	3,55%
CITROEN	185	2.267.083	12.255	98.481	305	2,49%
MERCEDES BENZ	40	1.865.629	41.641	95.532	1.388	3,33%

Fecha de última actualización: 13 de diciembre de 2008 00:03:38

Figura 4.18 Tasa promedio por marca de vehículo para el año 2007

InfoView - Microsoft Internet Explorer

Archivo Edición Ver Favoritos Herramientas Ayuda

Dirección http://localhost:8181/businessobjects/enterprise115/desktoplaunch/InfoView/main/main.do?objId=0

Analisis de Tasas

2007 CHEVROLET

TASA POR MARCA 2.007 - CHEVROLET

	Nro. Vhs Aseg.	Valor Asegurado	Valor Aseg. Prom.	Prima Cobrada	Prima Promedio	Tasa
LIUV	1297	25.787.743	19.867	639.117	493	2,48%
RODEO	53	546.654	10.314	631.218	11.810	116,47%
OPTRA	1448	21.928.536	15.144	622.947	430	2,84%
AVEO	1136	12.552.074	11.049	515.649	454	4,11%
CORSA	1829	16.801.749	9.196	462.198	253	2,75%
GRAND VITARA	333	4.590.590	13.796	147.922	444	3,23%
SPARK	520	8.133.396	15.641	83.930	161	1,03%
SUPER CARRY	91	943.458	10.369	30.096	331	3,19%
TRAILBLAZER	25	488.639	19.546	16.431	657	3,36%
CAMIÓN	2	11.000	5.500	550	275	6,00%
TOTAL =	6734	91.763.838	13.631	3.150.047	468	3,43%

Fecha de última actualización: 13 de diciembre de 2008 00:03:38

Figura 4.19 Tasa promedio para la marca CHEVROLET en el año 2007 profundizando en todos sus modelos (DRILL-DOWN)

4.2.4 REPORTE DE ANÁLISIS DE RECLAMOS

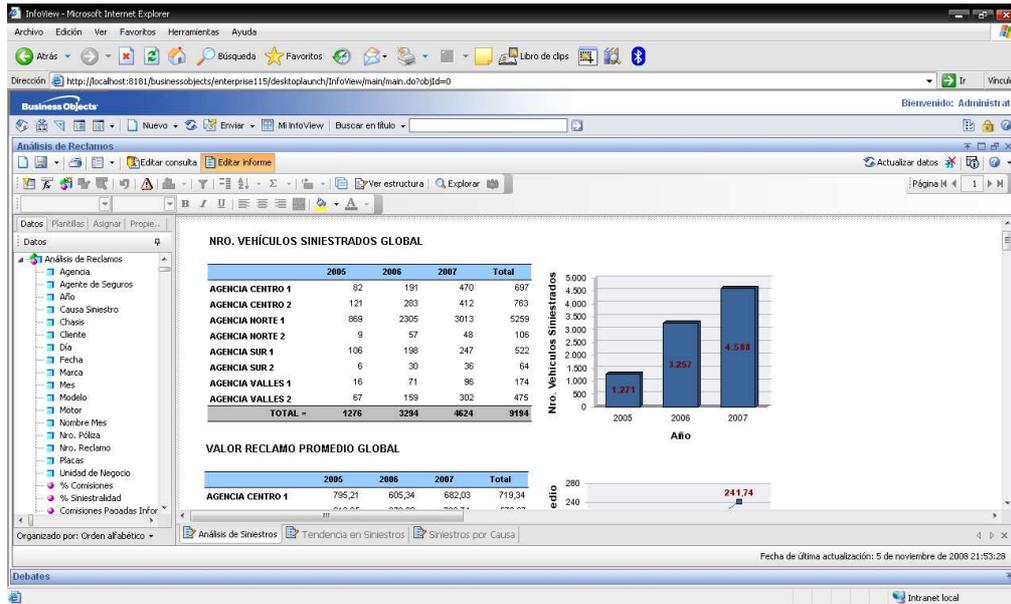


Figura 4.20 Pantalla de acceso al reporte de análisis de reclamos

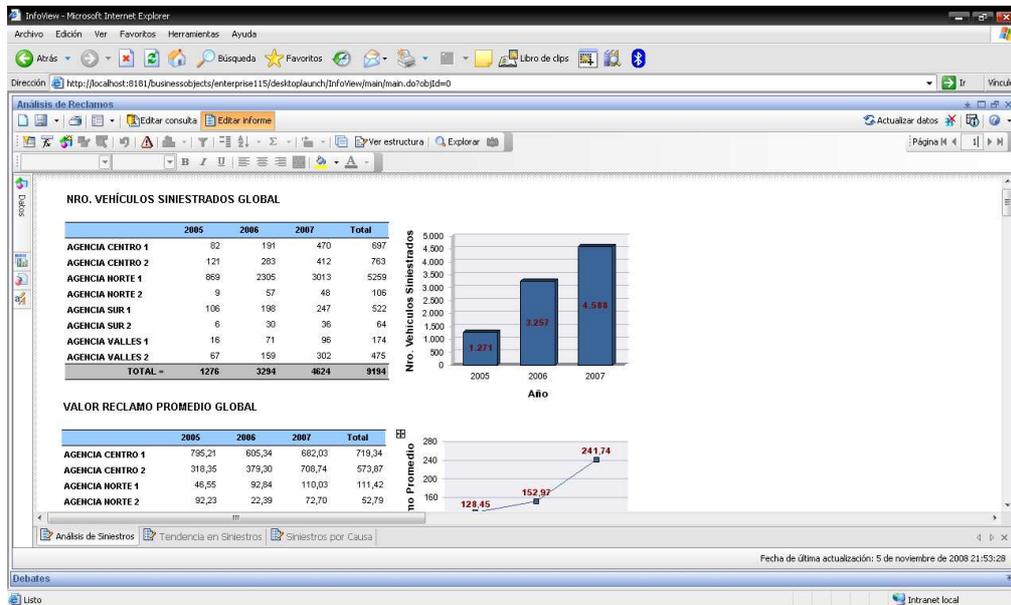


Figura 4.21 Número de vehículos siniestrados por agencia y año

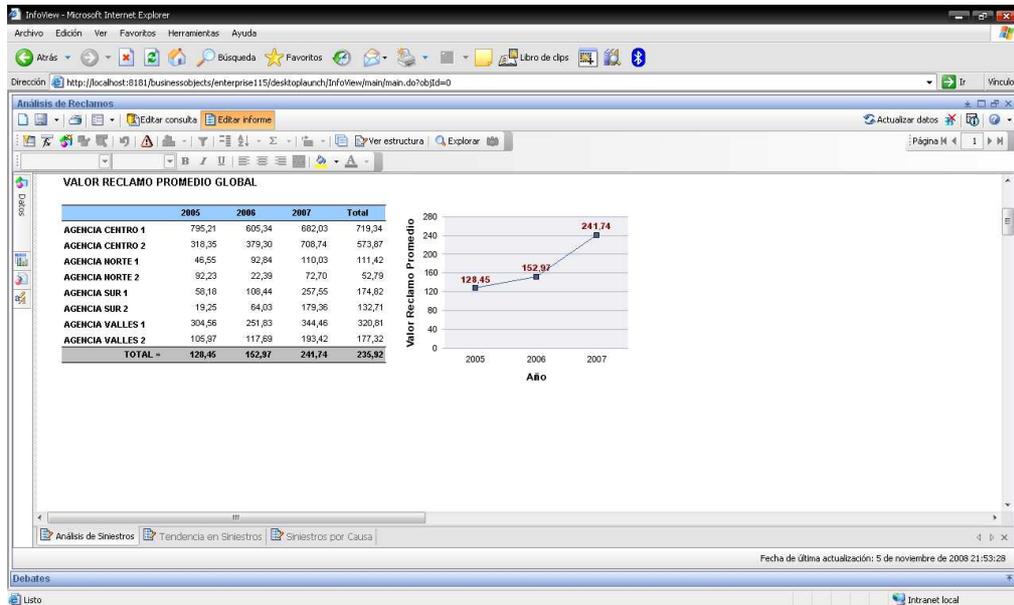


Figura 4.22 Valor reclamo promedio por agencia y año

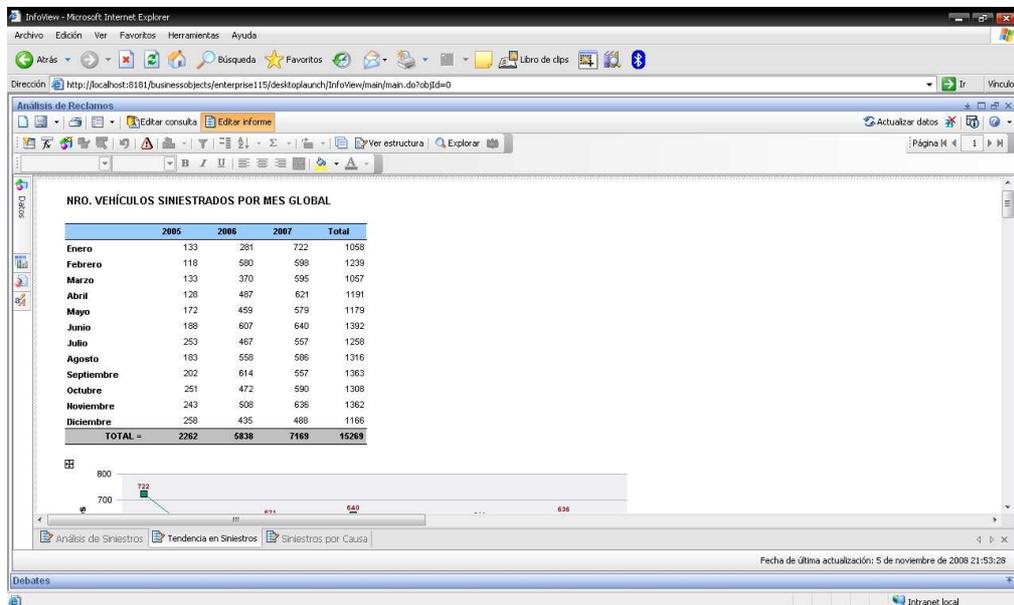


Figura 4.23 Número de vehículos siniestrados por mes y año de forma tabular

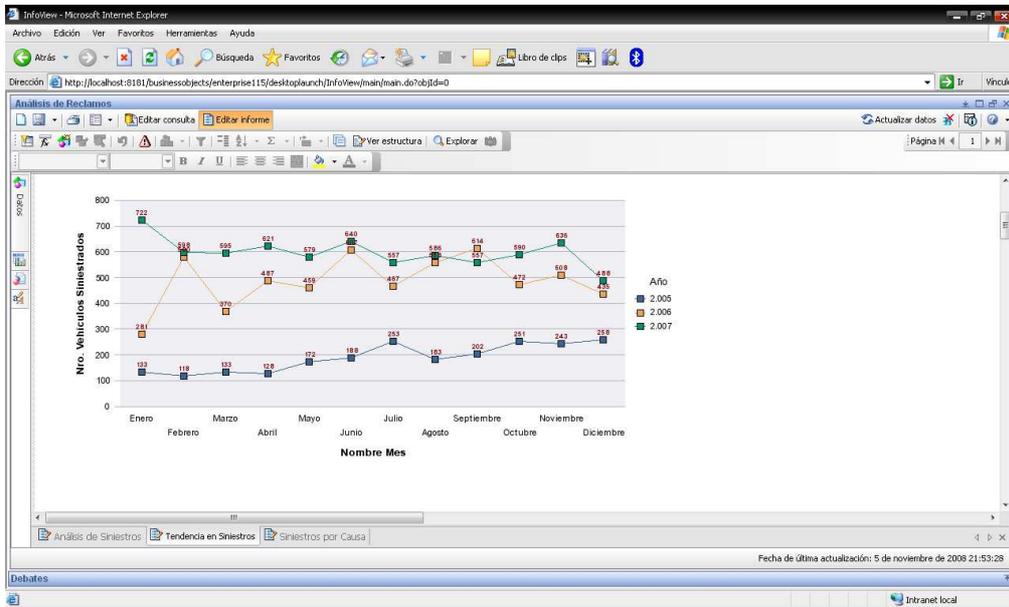


Figura 4.24 Número de vehículos siniestrados por mes y año en forma de gráfica de tendencias

	2005	2006	2007	Total
CHOQUE PARCIAL	994	2679	3759	6264
ROBO PARCIAL	176	327	463	890
CHOQUE TOTAL	31	73	123	194
CHOQUE Y VOLCADURA	24	69	77	149
RESPONSABILIDAD CIVIL	14	44	44	98
ROTURA DE VIDRIOS	14	28	42	80
OTROS	10	29	35	72
ROBO TOTAL	7	15	37	52
VOLCADURA	5	9	8	17
GASTOS MEDICOS	0	7	4	11
RESPONSABILIDAD CIVIL DAÑOS MATERIALES	0	5	4	9
DAÑO A TERCERAS PERSONAS	0	0	8	8
INCENDIO	1	3	2	5
RESPONSABILIDAD CIVIL DAÑOS CORPORALES	0	1	1	2
CATÁSTROFE	0	0	1	1
GASTOS MEDICOS A OCUPANTES	1	1	0	1
GASTOS MEDICOS A TERCERAS PERSONAS	0	0	1	1
TOTAL =	4277	3289	4689	9175

Figura 4.25 Número de vehículos siniestrados por causa y año de forma tabular

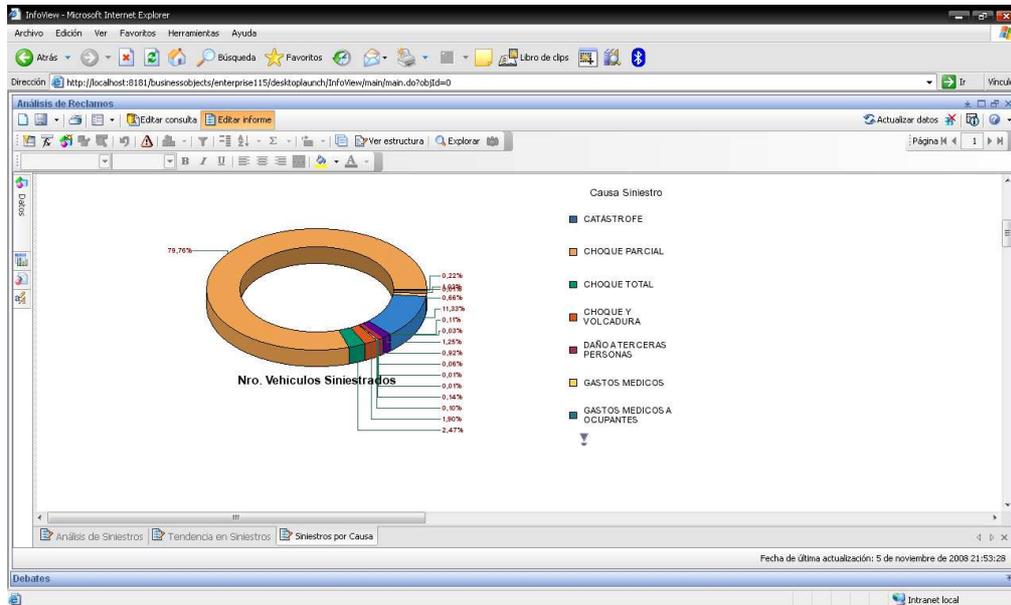


Figura 4.26 Porcentaje del número de vehículos siniestrados por causa para todos los años (2005 – 2007) en forma de gráfica de pastel

4.3 TABLERO DE MANDO INTEGRAL CON LOS INDICADORES DE GESTIÓN MÁS RELEVANTES

Para la elaboración del tablero de mando se utilizó la herramienta de gestión del rendimiento BusinessObjects Performance Management. Un tablero de mando está conformado por analíticas y métricas. Una métrica es un indicador de rendimiento real de la Compañía o KPI analizado en el tiempo. Una métrica puede ser comparada con un objetivo empresarial y evaluado con una tendencia.

Una analítica es la representación visual de una o más métricas. En la herramienta de gestión del rendimiento se puede crear las siguientes analíticas:

- Tendencias de métricas interactivas,
- Gráfico de Pareto,
- Velocímetro,
- Termómetro,
- Barómetro, y
- Semáforos

Las métricas utilizan un calendario para calcularlas. El calendario puede ser diario, mensual, trimestral, semestral, anual, o como requiera analizar una Compañía aseguradora. En el tablero de mando construido se utilizó analíticas con métricas mensuales.

El acceso al tablero de mando se lo realiza a través del portal Web de entrega de información InfoView accediendo a la opción de Gestión del Rendimiento como se muestra en la Figura 4.27.

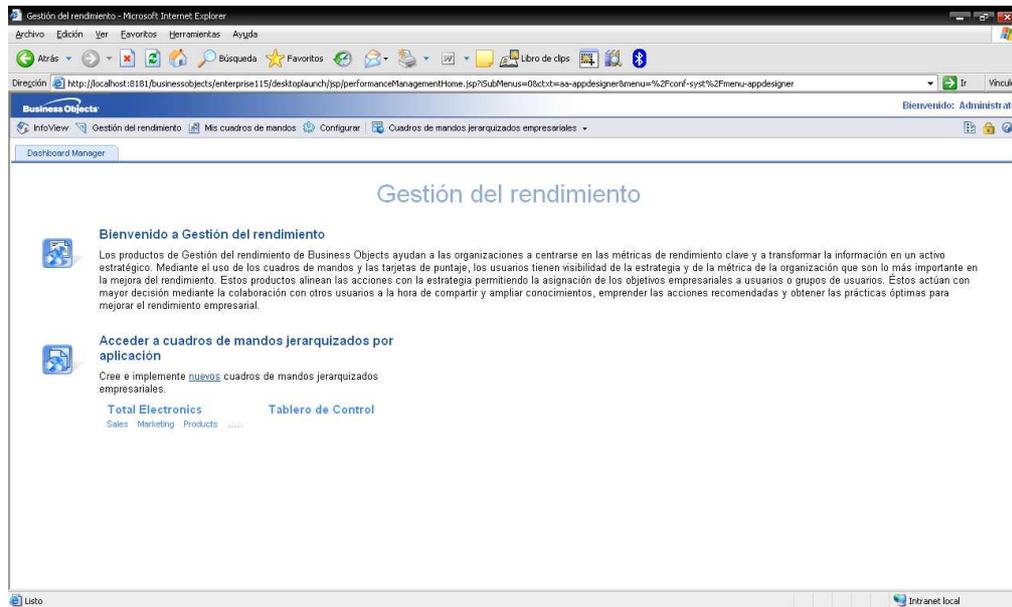


Figura 4.27 Pantalla de acceso al Tablero de Control de Seguros

El tablero construido se lo nombró como Tablero de Control, para acceder a éste es necesario hacer clic sobre el nombre del tablero. El tablero construido contiene cinco analíticas:

- Tendencia interactiva con el Resultado Técnico por Agencia,
- Barómetro con el porcentaje de siniestralidad por Agencia,
- Barómetro con la tasa promedio por Agencia,
- Barómetro con el porcentaje de comisiones por Agente de Seguros, y
- Un gráfico de Pareto con el Resultado Técnico por marca de vehículo.

En la Figura 4.28 se puede observar el tablero de control construido y las analíticas que se utilizó.

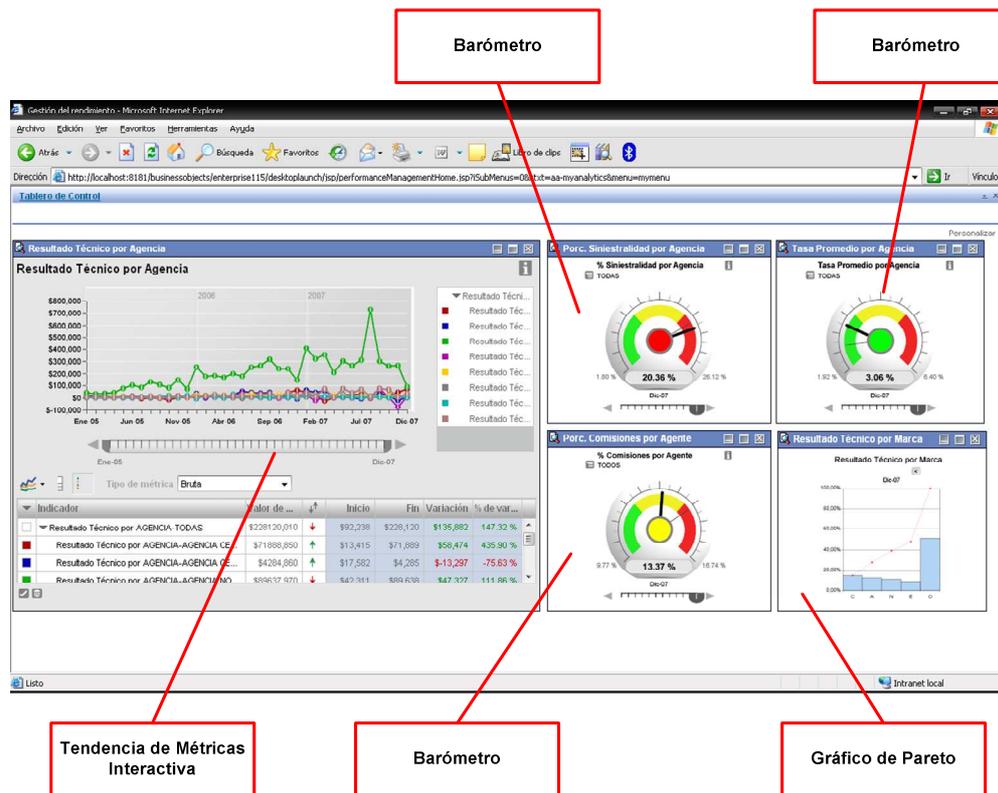


Figura 4.28 Tablero de control y analíticas utilizadas

A través de la tendencia de métricas interactiva de resultado técnico por agencia se puede analizar el valor del resultado técnico de cada agencia para los períodos que van desde Enero 2005 hasta Diciembre 2007. Es posible seleccionar ciertos períodos y ciertas agencias para realizar un análisis comparativo como se muestra en la Figura 4.29 en dónde se seleccionó el último semestre del año 2007 y se comparan las agencias CENTRO 1 y NORTE 1.



Figura 4.29 Tendencia de métricas interactiva para resultado técnico por agencia

En el barómetro del porcentaje de siniestralidad por agencia se puede analizar cuál ha sido la siniestralidad por agencia mensual desde los períodos Enero 2005 a Diciembre 2007. La analítica permite filtrar cierta agencia y seleccionar un período específico como se muestra en la Figura 4.30.

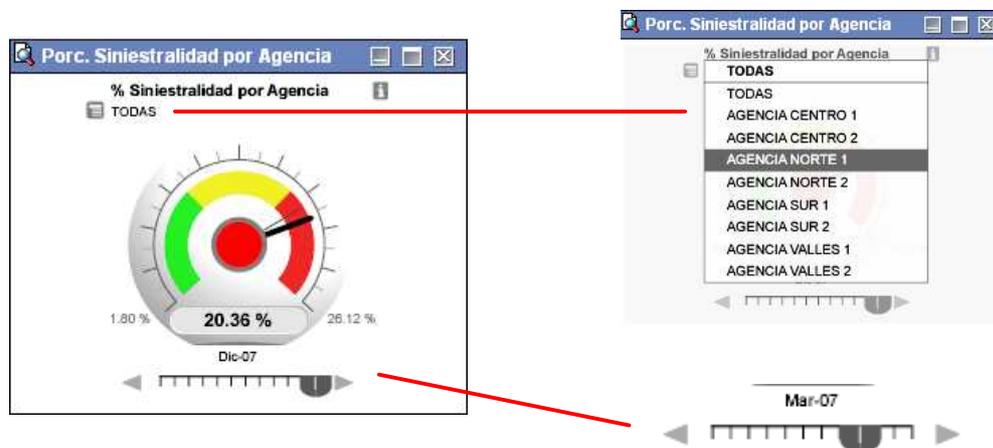


Figura 4.30 Barómetro del porcentaje de siniestralidad por agencia

El barómetro de la tasa promedio por agencia indica el valor de la tasa mensual con la que se han emitido las pólizas de vehículos por agencia y para los períodos comprendidos entre Enero 2005 y Diciembre 2007. En la Figura 4.31 se muestra la analítica.

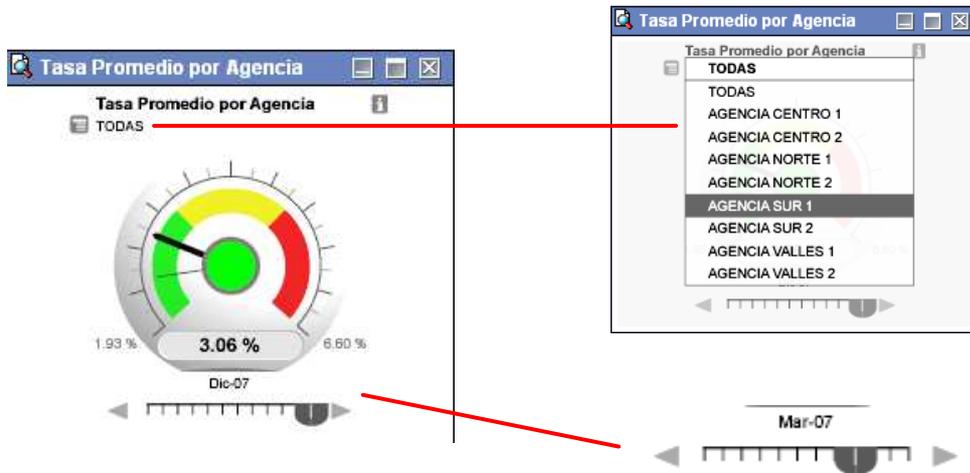


Figura 4.31 Barómetro de tasa promedio por agencia

En el barómetro del porcentaje de comisiones por Agente de seguros se puede analizar qué porcentaje de las primas se han pagado por concepto de comisiones a cada agente de seguros. La analítica puede ser analizada por agente de seguros y para cada período comprendido entre Enero 2005 y Diciembre 2007 como se muestra en la Figura 4.32.

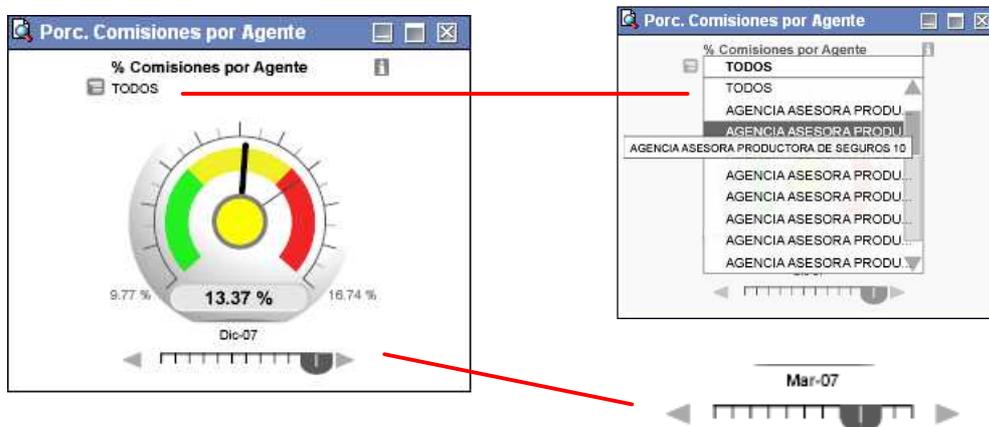


Figura 4.32 Barómetro del porcentaje de comisiones pagadas por agente de seguros

El gráfico de Pareto representa las cinco mejores marcas con mejor resultado técnico para los períodos comprendidos desde Enero 2005 hasta Diciembre 2007 como se puede observar en la Figura 4.33.



Figura 4.33 Gráfico de Pareto para el resultado técnico por Marca

CAPÍTULO V

5 CONCLUSIONES Y RECOMENDACIONES

5.1 CONCLUSIONES

- La inteligencia de negocios a través de actividades analíticas ayuda a una Compañía ha conocer en dónde ha estado, en donde está y en donde estará. Estas actividades analíticas realizadas a través de reportes, análisis OLAP o minería de datos apoyan al proceso de toma de decisiones en los niveles estratégicos, tácticos y operativos.
- Antes de comenzar un proyecto de inteligencia de negocios, y la construcción de un Data Warehouse es importante seleccionar y evaluar una arquitectura de inteligencia de negocios acorde al presupuesto y recursos de cada Compañía. Como se menciona en el Capítulo I, existen dos arquitecturas de inteligencia de negocios bien marcadas. La arquitectura propuesta por Bill Inmon y la arquitectura propuesta por Ralph Kimball, ambas tienen el mismo propósito pero tienen variantes significativas que pueden afectar directa o indirectamente al presupuesto y recursos de un proyecto.
- Muchas herramientas de inteligencia de negocios cumplen funciones específicas que apoyan a una actividad en especial en el desarrollo de un sistema de Data Warehouse. Es importante al momento de seleccionar las herramientas que todas se encuentren bajo una misma plataforma lo que permitirá lograr una estandarización de herramientas de BI. Es decir, todo el proceso de carga del sistema de Data Warehouse que va desde la obtención de datos de entrada hasta la presentación de información de salida estará soportada por herramientas que interactúan entre sí sin ningún inconveniente y cumplen con sus tareas específicas, evitándose problemas de

incompatibilidad e integración. En el proyecto de titulación realizado se seleccionó la plataforma de inteligencia de negocios BusinessObjects Enterprise XI Release 2 mostrándose el desarrollo del Data Mart para el análisis del ramo de vehículos apoyado por cada una de las herramientas de la plataforma, optimizándose tiempo y lográndose tener todo el sistema de Data Warehouse bajo una misma plataforma integrada.

- El desarrollo de un sistema de Data Warehouse es iterativo por naturaleza debido a la dinámica de un negocio, por ésta razón nace la premisa de que la metodología de desarrollo de un sistema de Data Warehouse tiene que ser iterativa. El sistema de Data Warehouse tiene que ser lo suficientemente flexible para responder a preguntas del negocio conforme vayan dándose los giros y cambios del negocio en el transcurso del tiempo.
- Es importante conocer los términos del negocio y los indicadores de gestión que apoyan al proceso de toma de decisiones. En el Capítulo II se hace énfasis en explicar de forma muy breve en qué consiste el negocio de seguros, cuáles son los términos del negocio, y cuáles son los indicadores que permiten evaluar el rendimiento del negocio asegurador.
- Resulta útil y productivo trabajar con una herramienta de integración de datos para implementar los procesos ETL de carga de un Data Mart. La herramienta de integración de datos permite consolidar datos de diversas fuentes, implementar los procesos ETL sobre éstos, y realizar las pruebas de carga necesarias hasta liberar los procesos de carga a un ambiente de producción.
- El definir una arquitectura de flujo de datos permite tener esquematizado cómo van a ir fluyendo los datos durante su transporte desde los sistemas fuente hasta el sistema de Data Warehouse. La arquitectura de flujo de datos constituye un mapa en dónde se podrá apreciar en qué base de datos se depositarán que datos durante todo su proceso de extracción, transformación y carga. Además, qué actividades se realizarán durante cada etapa del flujo de datos.

- Las herramientas de entrega de información a los usuarios finales tienen que ser lo suficientemente fáciles de usar para que ellos puedan realizar sus actividades analíticas en base a preguntas del negocio ad hoc. Es importante que la herramienta de análisis presente a los usuarios finales las dimensiones e indicadores de análisis en forma de piezas, las mismas que los usuarios podrán juntar y armar rompecabezas que resuelvan preguntas del negocio.
- La utilización de herramientas para cada etapa del proceso de construcción del Data Mart para el análisis del ramo de vehículos ha permitido ganar tiempo y tener todo integrado bajo una misma plataforma, la misma que puede ser modificada y actualizada de forma orquestada.
- El Data Mart para el análisis del ramo de vehículos para una empresa aseguradora muestra las dimensiones de análisis comunes en el negocio asegurador y los indicadores de análisis para evaluar el ramo de vehículos. Los indicadores de análisis de producción y siniestros ayudan a conocer el resultado técnico del ramo de forma histórica apoyando a la toma de decisiones del negocio asegurador.

5.2 RECOMENDACIONES

- Se recomienda utilizar herramientas que apoyen a cada etapa del proceso de construcción de un sistema de Data Warehouse. La utilización de herramientas permitirá optimizar el tiempo y controlar de mejor manera cada etapa. De igual forma que mejor que todas las herramientas sean parte de una única plataforma de inteligencia de negocios.
- Para un ambiente de producción se recomienda utilizar una metodología de desarrollo iterativa no en cascada como se utilizó en el desarrollo del presente proyecto de titulación debido a que el negocio es muy variable y el sistema de Data Warehouse tiene que ser lo suficientemente flexible para responder a las preguntas del negocio conforme el negocio vaya evolucionando.
- Un proyecto de inteligencia de negocios tiene que tener el patrocinio de la parte ejecutiva. Además, se recomienda que los objetivos empresariales

siempre estén alineados con los objetivos del proyecto de inteligencia de negocios, lográndose generar el valor del negocio necesario para mejorar la rentabilidad, disminuir costos y mejorar los procesos empresariales. Un proyecto de inteligencia de negocios no puede caminar sólo sin el apoyo y retroalimentación de la parte ejecutiva.

- El Data Mart que se ha construido en el presente proyecto de titulación puede ser utilizado como base para construir otro Data Mart de análisis para otros ramos generales del negocio asegurador. Las dimensiones e indicadores de análisis son comunes para el negocio asegurador a excepción de la dimensión vehículo que podría ser reemplazada por una dimensión predio, inmueble, etc. dependiendo del análisis del objeto asegurado que se desee realizar.

- La información que se puede generar a través del Data Mart construido en el proyecto de titulación podría ser utilizada como base para ampliar el alcance del proyecto, realizando actividades de minería de datos que permitan establecer tasas de seguros en base al análisis de grupos comunes como por ejemplo marcas, modelos, color de vehículo. El proceso de fijación de tasas de seguros de vehículos a partir de actividades de minería de datos en base a los datos históricos del Data Mart podría resultar en un proyecto de titulación interesante para futuros ingenieros.

BIBLIOGRAFÍA

- [1] BERNDTSSON, Mikael; HANSSON, Jörgen; OLSSON, Björn; LUNDELL, Björn. **Thesis Projects. A Guide for Students in Computer Science and Information Systems, Second Edition.** Segunda Edición. Springer-Verlag. 2008.
- [2] DAVIS, Jim; MILLER, Gloria; RUSSELL, Allan. **Information Revolution Using the Information Evolution Model to Grow Your Business.** Primera Edición. John Wiley & Sons, Inc. 2006.
- [3] HOBBS, Lilian; HILLSON, Susan; LAWANDE, Shilpa; SMITH, Pete. **Oracle Database 10g Data Warehousing.** Primera Edición. Elsevier Digital Press. 2005.
- [4] IMHOFF, Claudia; GALEMMO, Nicholas; GEIGER, Jonathan G. **Mastering Data Warehouse Design Relational and Dimensional Techniques.** Primera Edición. Wiley Publishing, Inc. 2003.
- [5] RAINARDI, Vincent. **Building a Data Warehouse: With Examples in SQL Server.** Primera Edición. Apress. 2008.
- [6] MAPFRE, Fundación. **Curso de Introducción al Seguro.** Segunda Edición. Fundación MAPFRE Estudios. 2001.
- [7] CLARK, P. **Business Objects Products and Services: the Industry's Leading Suite of Integrated Business Intelligence Products.** Business Objects. 2004.
- [8] KIMBALL, Ralph. **The data warehouse toolkit: the complete guide to dimensional modeling.** Segunda Edición. John Wiley & Sons, Inc. 2002.
- [9] WIKIPEDIA, The Free Encyclopedia. **Business Intelligence 2.0.** http://en.wikipedia.org/wiki/Business_Intelligence_2.0. Última visita, mayo 2008.
- [10] BusinessObjects, an SAP Company. **BusinessObjects XI 3.0:** <http://www.businessobjects.com/product/catalog/xi/> Última visita, mayo 2008.
- [11] Business Intelligence and Performance Management products from Cognos, an IBM Company. <http://www.cognos.com/> Última visita, mayo 2008.

- [12] Oracle and Hyperion. <http://www.oracle.com/hyperion/index.html>. Última visita, mayo 2008.
- [13] SAS, Business Intelligence Software and Predictive Analysis. <http://www.sas.com/> Última visita, mayo 2008.
- [14] Microsoft SQL Server: Analysis Services. <http://www.microsoft.com/sql/technologies/analysis/default.mspx> Última visita, mayo 2008.
- [15] Pentaho Comercial Open Source Business Intelligence. <http://www.pentaho.com/> Última visita, mayo 2008.
- [16] Eclipse BIRT Home. <http://www.eclipse.org/birt/phoenix/> Última visita, mayo 2008.
- [17] The Balanced Scorecard Institute. <http://www.balancedscorecard.org/>. Última visita, mayo 2008.
- [18] Intelligent Solutions, Inc. <http://www.intelsols.com/> Última visita, mayo 2008.
- [19] WIKIPEDIA, The Free Encyclopedia. **CODASYL**. <http://es.wikipedia.org/wiki/Codasy!> Última visita, mayo 2008.
- [20] WIKIPEDIA, The Free Encyclopedia. **Edgar F. Codd**. http://en.wikipedia.org/wiki/Edgar_F._Codd Última visita, mayo 2008.
- [21] WIKIPEDIA, The Free Encyclopedia. **Oracle Corporation** http://en.wikipedia.org/wiki/Oracle_Corporation Última visita, mayo 2008.
- [22] WIKIPEDIA, The Free Encyclopedia. **Bill Inmon**. http://en.wikipedia.org/wiki/Bill_Inmon Última visita, mayo 2008.
- [23] WIKIPEDIA, The Free Encyclopedia. **Ralph Kimball**. http://en.wikipedia.org/wiki/Ralph_Kimball Última visita, mayo 2008.
- [24] BusinessObjects Data Integrator. <http://www.businessobjects.com/product/catalog/dataintegrator/> Última visita, mayo 2008.
- [25] Oracle Database 10g Asynchronous Change Data Capture. http://www.oracle.com/technology/products/bi/db/10g/pdf/twp_cdc_cookbook_0206.pdf Última visita, mayo 2008.
- [26] WIKIPEDIA, The Free Encyclopedia. **Rate of return**. http://en.wikipedia.org/wiki/Return_on_investment Última visita, mayo 2008.
- [27] WIKIPEDIA, The Free Encyclopedia. **CRM**. http://en.wikipedia.org/wiki/Customer_relationship_management Última visita, mayo 2008.

- [28] WIKIPEDIA, The Free Encyclopedia. **Systems Development Life Cycle**.
http://en.wikipedia.org/wiki/Systems_Development_Life_Cycle Última visita,
junio 2008.
- [29] Seguro Obligatorio de Accidentes de Tránsito, SOAT Ecuador. **Normativa vigente relacionada con el SOAT**.
http://www.soatecuador.info/documentos/Decreto_809_Soat.pdf Última visita,
julio 2008.
- [30] MEYTHALER, Eduardo. **El Contrato Partes y Elementos –Aspectos Legales**. Instituto de Prácticas Bancarias y Financieras (IPBF), Programa de Especialización en Seguros. 2000.
- [31] TARRÉ, Julio. **Seguros Generales – Vehículos**. Instituto de Prácticas Bancarias y Financieras (IPBF), Programa de Especialización en Seguros. 2000.
- [32] Superintendencia de Bancos y Seguros – Ecuador. Sistema de Seguros Privados – Ley General de Seguros. **Ley General de Seguros**.
http://www.superban.gov.ec/downloads/normativa/ley_gral_seguros.pdf Última visita, agosto 2008.
- [33] Superintendencia de Bancos y Seguros – Ecuador. Sistema de Seguros Privados – Ley General de Seguros. **Legislación sobre el Contrato de Seguros - Decreto Supremo 1147**.
http://www.superban.gov.ec/downloads/normativa/decreto_supremo_1147.pdf
Última visita, agosto 2008.
- [34] LEXIS Ecuador. Legislación Indexada Sistemática. **Codificación del Código Civil de la República del Ecuador**.
<http://www.lexis.com.ec/lexis/novedadesDescargas/CodigosLeyes/CODIGO%20CIVIL.htm> Última visita, agosto 2008.
- [35] Business Objects. **Data Integrator Getting Started Guide**. Febrero, 2006.

ANEXOS

ANEXO A. Instalación y Configuración del servidor de base de datos Oracle 10g

ANEXO B. Instalación y Configuración de BusinessObjects Data Integrator

ANEXO C. Creación de la base de datos dimensional (DIM_DATA)

ANEXO D. Instalación y Configuración de BusinessObjects Enterprise XI Release 2 Service Pack 3

ANEXO E. Metadatos del universo DM_VEHIC

ANEXO F. Metadatos del universo DM_VEHIC_PM

ANEXO G. Diccionario de datos, base de datos de Seguros de Vehículos